(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

47826

## The General Field of Psychology

EDITED BY

CARL MURCHISON

Gordon Allport
Harvard University

Arthur G. Bills
University of Cincinnati

E. A. Bott
University of Toronto

Leonard Carmichael
Smithsonian Institution

Herbert S. Conrad
U. S. Office of Education

Elmer K. Culler
University of Rochester

John F. Dashiell
University of North Carolina

Roy M. Dorcus
University of California at Los Angeles

Horace B. English
Ohio State University

Franklin Fearing
University of California at Los Angeles

Frank A. Geldard
University of Virginia

Clarence H. Graham
Columbia University

Ernest R. Hilgard
Stanford University

H. M. Johnson
Tulane University

Heinrich Klüver
University of Chicago

Carney Landis
Columbia University

K. S. Lashley
Harvard University

Norman L. Munn
Bowdoin College

John Paul Nafe
Washington University

Donald G. Paterson
University of Minnesota

Lee Edward Travis
University of Southern California

Robert C. Tryon
University of California

Michael J. Zigler
Wellesley College

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 37+38

1954

Copyright, 1954, by The Journal Press

Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

LIBRARY

# AUTHOR INDEX

Amatora, Sister Mary . . . 45
Andreas, B. G. . . . 173, 185
Auble, D. . . . 251
Barrera, Elizabeth . . . 199
Bingham, W. E., Jr. . . . 225
Bradbard, L. . . . 27
Briggs, Dennie . . . 233
Coleman, James C. . . . 243
Corso, John S. . . . 297
Culbert, Sidney S. . . . 81
Dana, Richard H. . . . 139
Essenberg, J. M. . . . 291
Friedman, Merton H. . . . 133
Fuller, Nuni . . . 65
Glaser, Nathan M. . . . 95
Gottsdanker, Robert M. . . . 123
Green, R. F. . . . 173, 185
Haslerud, G. M. . . . 27
Hellman, Leon I. . . . 243
Hirsch, Richard S. . . . 257
Holzman, Philip S. . . . 105
Hurst, F. . . . 251
Janis, Irving L. . . . 75
Kapos, E. . . . 251
Johnson, R. P. . . . 27
Klein, George S. . . . 105
Klüver, Heinrich . . . 199
Knehr, Charles A. . . . 65
Landis, Carney . . . 65
Lepley, William M. . . . 19
Mayzner, Mark S., Jr. . . . 135
Mech, E. . . . 251
Meer, Bernard . . . 39
Milam, Albert T. . . . 31
Milholland, Harry C., Jr. . . . 75
Neff, Walter S. . . . 95
Norris, Eugenia B. . . . 163
Plutchik, Robert . . . 59
Ross, Bruce M. . . . 279
Rubin-Rabson, Grace . . . 151
Saul, Ezra V. . . . 257
Schneiderman, Leo . . . 155, 271
Spragg, S. D. S. . . . 163, 173, 185
Stein, Morris I. . . . 39
Sumner, F. C. . . . 31
Swartz, Paul . . . 163
Tresselt, M. . . . 135
Wedemeyer, Barbara . . . 51

Bureau Edn. Research
COLLEGE
11·8·70
No. g 112

## TABLE OF CONTENTS

Crozier and Wolf on flicker-fusion, 1933-1944 . . . . . . . . . . 3
By Carney Landis

Variability as a variable . . . . . . . . . . . . . 19
By William M. Lepley

Pure guidance and handling as components of the Maier technique for breaking abnormal fixations . . . . . . . . . . . . . 27
By G. M. Haslerud, L. Bradbard, and R. P. Johnson

Spread and intensity of vocational interests and evaluative attitudes in first-year negro medical students . . . . . . . . . . . 31
By Albert T. Milam and F. C. Sumner

Perceptual organization in a study of creativity . . . . . . . . 39
By Morris I. Stein and Bernard Meer

Similarity in teacher and pupil personality . . . . . . . . . 45
By Sister Mary Amatora

Rorschach statistics on a group of 136 normal men . . . . . . . 51
By Barbara Wedemeyer

Further remarks on the hypothetical construct . . . . . . . . 59
By Robert Plutchik

Sensory vs. autonomous control of span of apprehension . . . . . 65
By Charles A. Knehr and Nuni Fuller

The influence of threat appeals on selective learning of the content of a persuasive communication . . . . . . . . . . . 75
By Irving L. Janis and Harry C. Milholland, Jr.

Directional after-effects following systematic distortion of the visual field 81
By Sidney S. Culbert

Normative data on the Rorschach . . . . . . . . . . . . 95
By Walter S. Neff and Nathan M. Glaser

Cognitive system-principles of leveling and sharpening: Individual differences in assimilation effects in visual time-error . . . . . . . . 105
By Philip S. Holzman and George S. Klein

The continuation of tapping sequences . . . . . . . . . . 123
By Robert M. Gottsdanker

Two-factor interpretation of an instance of fear learning . . . . 133
By Merton H. Friedman

A further experiment in the recognition of ego-involved materials . . . 135
By M. E. Tresselt and Mark S. Mayzner, Jr.

Personality orientation: An organizational focus for current research . . 139
By Richard H. Dana

Intelligence and conservative-liberal attitudes . . . . . . . . 151
By Grace Rubin-Rabson

Social perception as a function of identification . . . . . . . 155
By Leo Schneiderman

Performance on a following tracking task (modified SAM two-hand coördination test) as a function of radius of control cranks . . . . 163
By Paul Swartz, Eugenia B. Norris, and S. D. S. Spragg

Transfer effects between performance on a following tracking task (modified SAM Two-Hand Coördination Test) and a compensatory tracking task (modified SAM Two-Hand Pursuit Test) . . . . . . . . 173
By B. G. Andreas, R. F. Green, and S. D. S. Spragg

Transfer effects in following tracking (modified SAM Two-Hand Coördination Test) as a function of reversal of the display-control relationships on alternate blocks of trials . . . . . . . . . . . . . 185
By B. G. Andreas, R. F. Green, and S. D. S. Spragg

On the use of azaporphin derivatives (phthalocyanines) in staining nervous tissue . . . . . . . . . . . . . . . . . 199
By Heinrich Klüver and Elizabeth Barrera

Electromagnetic and electrostatic fields: A neglected area in physiological psychology . . . . . . . . . . . . . 225
By W. E. Bingham, Jr.

A modification of the Thematic Apperception Test for Naval enlisted personnel (N-TAT) . . . . . . . . . . . . . . 233
By Dennie L. Briggs

An experimental evaluation of the relative effectiveness of insulin coma and electronarcosis therapies with schizophrenic patients: I. Two weeks after treatment . . . . . . . . . . . . . . . . . 243
By James C. Coleman and Leon I. Hellman

A mathematical description of the relationship between frequency of responding (F) and intertrial interval in a motivational situation . . . 251
By E. Mech, E. Kapos, F. Hurst, and D. Auble

Psychological problems in marksmanship of infantry type weapons . . . 257
By Ezra V. Saul and Richard S. Hirsch

Anxiety and social sensitivity . . . . . . . . . . . 271
By Leo Schneiderman

Probability predictions between and within series . . . . . . . 279
By Bruce M. Ross

The effect of nicotine on maze behavior of albino rats . . . . . . 291
By J. M. Essenberg

Scale position and performed melodic octaves . . . . . . . . 297
By John F. Corso

$7.00 per volume
$4.00 per half volume

QUARTERLY
Two volumes per year
Immediate publication

January, 1954
Volume 37, First Half

(Founded in 1935 by Carl Murchison)

478 Z 7

# The Journal of PSYCHOLOGY

## The General Field of Psychology

---

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

## JANUARY, 1954

**(Manuscripts are printed in the order of final acceptance)**

Crozier and Wolf on flicker-fusion, 1933-1944 . . . . . . . . . 3
By Carney Landis

Variability as a variable . . . . . . . . . . . . 19
By William M. Lepley

Pure guidance and handling as components of the Maier technique for breaking abnormal fixations . . . . . . . . . . . . 27
By G. M. Haslerud, L. Bradbard, and R. P. Johnson

Spread and intensity of vocational interests and evaluative attitudes in first-year negro medical students . . . . . . . . . . . 31
By Albert T. Milam and F. C. Sumner

Perceptual organization in a study of creativity . . . . . . . 39
By Morris I. Stein and Bernard Meer

Similarity in teacher and pupil personality . . . . . . . . 45
By Sister Mary Amatora

(over)

---

Copyright, 1954, by The Journal Press
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter January 26, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

Rorschach statistics on a group of 136 normal men . . . . . . . 51
By Barbara Wedemeyer

Further remarks on the hypothetical construct . . . . . . . . . 59
By Robert Plutchik

Sensory vs. autonomous control of span of apprehension . . . . . 65
By Charles A. Knehr and Nuni Fuller

The influence of threat appeals on selective learning of the content of a persuasive communication . . . . . . . . . . 75
By Irving L. Janis and Harry C. Milholland, Jr.

Directional after-effects following systematic distortion of the visual field 81
By Sidney S. Culbert

Normative data on the Rorschach . . . . . . . . . . . 95
By Walter S. Neff and Nathan M. Glaser

Cognitive system-principles of leveling and sharpening: Individual differences in assimilation effects in visual time-error . . . . . . . 105
By Philip S. Holzman and George S. Klein

The continuation of tapping sequences . . . . . . . . . . 123
By Robert M. Gottsdanker

Two-factor interpretation of an instance of fear learning . . . . 133
By Merton H. Friedman

A further experiment in the recognition of ego-involved materials . . . 135
By M. E. Tresselt and Mark S. Mayzner, Jr.

Personality orientation: An organizational focus for current research . . 139
By Richard H. Dana

Intelligence and conservative-liberal attitudes . . . . . . . . 151
By Grace Rubin-Rabson

Social perception as a function of identification . . . . . . . . 155
By Leo Schneiderman

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 3-17.

# CROZIER AND WOLF ON FLICKER-FUSION, 1933-1944*

*Psychiatric Institute, Columbia University*

---

CARNEY LANDIS[1]

---

478Z75

## A. INTRODUCTION

Crozier and Wolf and their collaborators (Zerrahn-Wolf, Holway) published a long series of papers dealing directly or indirectly with the problem of flicker-fusion between the years 1933 and 1944. They never published in an integrated form an over-all statement of their theory or findings. All but one of their publications were in either the *Journal of General Physiology* or in the *Proceedings of the National Academy of Science,* journals not usually consulted by psychologists or by many other investigators working in the field of vision. Their approach to the problems of flicker-fusion was different from that usually employed. Their published papers are not easy reading. In the entire series only occasional reference was made to the work of any other investigator except Hecht. With Hecht's theory and explanations they continued a running battle.

Crozier and Wolf obtained frequency-intensity contours on a greater range of animal species than anyone else before or since. As a general rule they give a very adequate statement of the conditions of their experiments and publish in full a tabulation of their obtained determinations. They limited their human studies for the most part to WJC's left eye though some data were obtained for his right eye and binocular regard and for EW. (They never compared their findings to the complete published data of other investigators for other human eyes.) With WJC's L eye they investigated the effects of more determinants and combinations of determinants than any previous or subsequent investigator, giving the relevant tabulation of determinations.

Based on a very complete survey of the publications dealing with flicker-fusion, one can state that in addition to the range of their investigations they made at least four original and unique findings. These were, (*a*) the rôle of genetics in the determination of the *F-log I* curve, (*b*) the effect

---

*Received in the Editorial Office on July 1, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This survey was facilitated by a grant from the Rockefeller Foundation. The author expresses his indebtedness to his colleagues and students who worked through this material with him in a seminar in 1952.

of a striated test field in contrast to a uniform field, (*c*) the avian pecten effect, and (*d*) their method for a computational basis of the comparison of the effect of the physical intensity of stimulation by monochromatic light of differing hues.

In the following itemized presentation I have selected direct statements or have paraphrased their published statements on points of relevance and importance. The particular reference from which the item is drawn is usually given, but in some instances the statement or idea is one drawn from several references especially when it was an idea which was modified from time to time as the research progressed. In this latter case specific reference is omitted.

## B. Theoretical Assumptions

1. In the study of many varieties of biological phenomena the indices of variability of measurement are as informative as are the means. Hence in the examination of the variability of measurement one must inspect every possible source and correctly apportion the total variability to the proper sources (1).

2. Only when all sources of variability have been examined and verified and the differences between two or more individuals shown to be inconsequential is it permissible to combine data obtained from different individuals and to present grouped data as homogeneous (12).

3. The threshold of excitation is the reciprocal of intensity in millilamberts. This threshold varies from neural unit to neural unit, and from time to time in the same unit, in the fashion of a probability distribution curve (15).

4. The sensation of brightness is a function of the total neural effect (integration of activity of all retinal, connecting and cortical neural units acting under given circumstances at a given time) (14).

5. Intensity discrimination ($\Delta I$) is the result of the change in units of neural effect from some existing intensity, either $I_0$ or $I_1$ (14).

6. The effect of any increase or decrease in stimulus intensity is largely determined by the difference in neural effect already existing and the maximum possible effect (14).

7. The assumption that visual data are determined only by quantitative properties of the retina is unnecessary and fallacious. The parallelism of known retinal properties derives simply from the fact that a large number of variable elements are involved in each of the two cases while the invariants are essentially statistical properties (13).

## C. NOTATION

$A$ = area of test patch.

$B$ = binocular.

$CFF$ = critical flicker-fusion threshold.

$D$ = dark interval.

$\Delta$ = discriminable difference; just noticeable difference.

$e$ = base for natural logarithms.

$F$ = frequency; rate of intermittence in cycles per second (*cps*).

*F-log I Curve* = graphic representation of relation of $F$ (ordinate) to $\log I$ (abscissa).

$F_m$ = mean frequency when $F$ is the dependent variable.

$F_{max}$ = highest possible rate of intermittence that can be obtained with maximum $I$ available; $F_{max}$ becomes asymptotic to an increasing $I$.

*gamma* = subjective pulsing of the stripes in a striated field at less than the rate of physical intermittence of illumination.

$I$ = intensity (usually millilamberts).

$I_m$ = mean intensity when $I$ is the dependent variable.

$K$ or $k$ = constants.

$L$ = left eye; light interval.

$$LDR = \frac{L}{L+D} = \frac{\text{Duration of flash in one cycle of repetitive intermittence } (L)}{L \text{ plus duration of Darkness in one cycle of repetitive intermittence}}$$

*phi* = subjective pulsing, side-to-side movement of test field.

$R$ = right eye; red.

$\sigma$ = standard deviation.

$\sigma'_{\log I}$ = the first derivative of *F-log I* regarded as a probability integral; it is inversely proportional to the slope of *F-log I* curve.

$\tau'$ (tau prime) = the inflection point of the *F-log I* curve.

$t_D$ = duration of dark period in one cycle of repetitive, intermittent illumination.

$t_L$ = duration of light flash in one cycle of repetitive, intermittent illumination.

$V$ = violet, $Bl$ = blue, $G$ = green, $Y$ = yellow, $R$ = red, $W$ = white.

## D. THE NATURE OF FLICKER-FUSION

1. Crozier and Wolf say theirs is the intensity discrimination theory of the marginal (critical) recognition of flicker. This implies that the effects discriminated are due to (*a*) the average action of a flash of light, and (*b*) those represented by the after-image or after-action persisting during the

dark interval (no light). Flicker recognition is a forced response due to summative action of the elements of effect produced in a population of intrinsically fluctuating units (6, 23).

2. Flicker-fusion thresholds occur when the relation between *F* and *I* is such that the average effect of a light flash (or the total effect during a light interval) just fails to be distinguishable from the effect due to or expressed as the after-image of the flash (3).

3. The elementary problem of visual flicker is to deduce from the properties of the data the nature of the system in which these properties are determined and by which their exhibition is governed, which is a problem of statistical mechanics (13).

4. The transition from fusion to flicker is a special case of intensity discrimination. The transition from flicker to fusion is a different process than vice versa, having a greater variability; it is not a simple intensity discrimination process (6).

## E. METHODS AND MATERIAL

1. Two methods were used in the study of flicker-fusion. The methods were a function of two kinds of apparatus; (*a*) a striped cylinder which revolved vertically around the organ of vision; (*b*) the discriminometer which provided two beams of parallel rays of light which fall on either or both eyes. This latter method was used only with human observers, namely Crozier and Wolf (2, 12).

2. In the striped cylinder method used with animal forms other than man the *CFF* threshold was indicated by some unique response of the animal to the apparent change of fusion to flicker; WJC and EW signaled their *CFF* thresholds verbally. *CFF* contours were obtained for some 30 species including the honeybee, dragonfly, a variety of species of small fish, several crustaceans, several varieties of lizards, salamanders, frogs, turtles, and birds (27).

3. In almost every study a table providing full data obtained is given. Their *I* varied between log —7.0 and log 3.5; *F* from 1 to 62 *cps*.

## F. DETERMINANTS

1. The fundamental functional activity investigated was intensity discrimination which may show itself in the form of visual acuity, course of adaptation, or flicker-fusion (12).

2. With relation to time the intermittent repetition may provide intervals of *I* which vary in duration, or they may be constant; presentation may be continuous and uniform; or abrupt, discrete, and single (12).

3. With respect to retinal area the determinants may be:
   - (*a*) Two simultaneous *I*'s.
   - (*b*) Two *I*'s in different retinal locations.
   - (*c*) Two *I*'s in succession.
   - (*d*) I = 0 versus *I* = *anything*.
   - (*e*) Two *I*'s of different hue.
   - (*f*) Two *I*'s of different duration of exposure (12).
4. The important flicker determinants are:
   - (*a*) Wave-length composition of the light which includes both *I* and hue.
   - (*b*) Form of the light-dark alternation.
   - (*c*) Position and size of the light stimulus patch on the retina.
   - (*d*) Constancy of the pupillary opening.
   - (*e*) State of adaptation.
   - (*f*) Length of exposure of eye to the test patch.
   - (*g*) Uniformity or striation of test patch.
   - (*h*) Light time fraction of the flash cycle.
   - (*i*) The temperature and homogeneous state of the organism (15).

5. Actually studied were the following determinants: (*a*) *F,* (*b*) *I,* (*c*) Geometric form of the eye, (*d*) *LDR,* (*e*) Area, (*f*) Position of *A* on retina, (*g*) Fovea v. periphery of retina, (*h*) Rod v. cone action, (*i*) Binocular v. monocular regard, (*j*) Temperature of organism, (*k*) Genetic constitution of organism, (*l*) Avian pecten, (*m*) Hue, (*n*) Effect of drugs on pupil size, (*o*) Uniformity v. striation of test area, (*p*) Age of observer, (*q*) Subjective *CFF* threshold points, viz., gamma and phi, (*r*) Wave form of stimulus flash, (*s*) Practice, (*t*) Interaction of two or more of the foregoing list.

## G. THE *F-log I* CONTOUR

1. For reasons never explicitly stated the form of the *F-log I* contour and the effect of one or several determinants on this contour constitutes the central point of the Crozier and Wolf studies.

2. Although Hecht and others have speculated otherwise, Crozier and Wolf are satisfied to repeatedly demonstrate *F-log I* is a simple probability integral so far as the "cone" section is concerned; the "rod" portion is likewise a probability integral curve having (usually) different values for its parameters.

3. The necessary parameters to describe the *F-log I* contour are $F_{max}$, $\tau'$ (tau prime) and $\sigma_{log I}$. $F_{max}$ is the highest rate of intermittence at which a flicker-fusion threshold can be found. (How it is computed, other than em-

pirically is never stated.) Tau prime is the inflection point of the *F-log I* contour; when plotted on a logistic grid it is the *log I* corresponding to that *F* which is 50 per cent of $P_{max}$. Sigma prime *log I* is inversely proportional to the slope of the curve; when plotted on a logistic grid it is the range of *log I* between $F\% = 50$ and $F\% = 84$ (7).

4. The *F-log* $I_m$ curve is significantly different in position from the $F_m$*-log I* curve. The difference rests on the fact that $\sigma_F$ and $\sigma_I$ represent different dispersions of measurements of properties of the reacting organism and are not due to experimental errors. The variation in $F_m$ corresponds to organic fluctuations of the organism, e.g., pulse, while variation in $I_m$ reflects variations in the potential of the chemical processes the velocity of which determines the interval between fixed levels of performance (6).

## H. THE SLOPE ($\sigma'_{log\ I}$) OF *F-log I* CONTOURS

1. For a given retinal region the slope is a function of the number of neural units which is determined as a genetic dominant (23).

2. In most situations or comparisons the slope gives more information than either $F_{max}$ or tau prime since the slope is slightly (if at all) affected by hue, *LDR,* temperature, retinal pigmentation, and gamma or phi. It is affected by subdivision of the image, striation, age, genetic constitution, geometry of the eye, i.e., flat v. curved surfaces, inhibitory effect of cones or rods, combining variation of hue and *LDR,* and $t_D$. Combining *LDR* and *F* while varying temperature frees the slope from $t_D$ effect (27).

## I. TAU PRIME OR THE INFLECTION POINT

1. Tau prime is the $I_m$ where $F/F_{max}$ is 50 per cent. It is the homologue of chronaxie (tau) in electrical stimulation experiments. It represents the point where 50 per cent of the total potential units of effect are activated (8).

2. Tau prime is independent of retinal area, retinal location, and temperature. It is altered rectilinearly with changing *LDR* or $t_D$, by changing hue (other than white), by striation particularly at *LDR* .75 and .90, and by genetic blending (24, 26, 27).

## J. $F_{max}$

1. $F_{max}$ represents the point where the units of effect are operating at their absolute limit of efficiency for the particular conditions of any experiment. All of the published tables give the highest *F* which was obtained together with its $I_m$ value. In all except one table (Ref. *27,* Table 1) the

$F_{max}$ computed for logistic grid plots is higher than the highest obtained $F$ (17).

2. At higher $I$ levels $F_{max}$ is not changed to any considerable extent by increasing $I$. However they never found a decline in $F$ with an increasing $I$ at higher $I$ levels as Hecht and others found (17).

3. With a uniform field $F_{max}$ is inversely related to *LDR;* with divided or striated field *LDR*'s over .75 reach higher $F_{max}$ than those of .50 and below (26).

4. $F_{max}$ is independent of temperature. Since it is usually found to have a frequency between 50 and 60 *cps* the general minimum time for optimum photochemical reaction is about 0.009 sec. at *LDR* .50 (4).

5. $F_{max}$ is dependent on *LDR,* hue, area, location, and form of area. Of all determinants it is independent only of temperature (23).

6. By adjusting *LDR* to give equal $F_{max}$'s the effect on slope and tau prime of any other independent variable may be found (20).

7. Usually Crozier and Wolf obtained the $F_{max}$ of the "rod" portion of curves by statistical dissection and extrapolation. In geckos with a pure rod retina *F-log I* is a simplex curve with $F_{max}$ of 26.5 and there is no difference in $F_{max}$, tau prime, or slope with either $F_m$ or $I_m$ (13A, 20).

### K. FOVEA V. PERIPHERY; CONES V. RODS; DUPLEXITY THEORY

1. The photochemistry of the rods and cones is not different since they respond in the same way to temperature changes but not to striations of the field. Hence differences found are statistical consequences of differences in excitation thresholds (27).

2. In a duplex retina, cones inhibit rod effects but rods do not alter cone effects (24).

3. Rods and cones are independently affected by hue, retinal location, subdivision of the field, and pecten effect. They are affected in the same way by *LDR,* temperature, and binocularity (27).

4. Suppose the "rod" portion of the *F-log I* curve is extrapolated backward to find its rate of decline. Then it may be shown that the more rapid the ascent of the "cone" portion the more abrupt the descent of the extrapolated "rod" portion. This is further evidence that flicker is not retinally but centrally determined (19).

5. Neither $F_{max}$, tau prime, nor slope is associated in a constant fashion with any determinant for both rods and cones (27).

6. Comparison of simplex curves of the turtle (cone retina) with the gecko (rod retina) gives no evidence that rods function best at low $I$ and cones at higher levels of $I$ (13A).

7. A real competitive integration of rod and cone effects occurs in the recognition of flicker so that the foveal attention of the observer determines the fact that the levels of foveal vision predominate. Both the subjective appearance of the field and the objective *CFF* threshold indicate that discrimination is a central nervous system and not a retinal function (19).

8. Sunfish "cone" curves show two temperature coefficient slopes with change in temperature; their "rod" curves show only one temperature coefficient (5).

9. The change of slope in the *F-log I* curve for Crozier's left eye over a 2½-year period was greater in the "rod" than in the "cone" portion (17).

10. The retinal duplex structure is associated with the duplex *F-log I* curve. However the "rod" or "cone" portion of the curve is not solely dependent on rods or cones. The rod effect and cone effect can be demonstrated in animals having no rods or no cones (27).

11. The hereditary determination of the rod and cone effects is different (13).

12. In man, violet involves the fewest cone units, red the most. The smaller number of cone units provides the largest number of elements of effect (20).

13. In Newt triturus the *F-log I* curve shows three portions; hence there must be a third retinal process here in addition to that of rods and cones (16).

### L. AREA; SIZE AND LOCATION

1. In man the basic form of the *F-log I* curve is slightly or not at all a function of area (8).

2. *F* increases with *A* at a given location. Within a limited range *A* can be substituted for *I*. Doubling *A* in case of monocular versus binocular regard, *F* (or *log I*) is increased in the proportion 1:1.4 (5, 17).

3. Reduction of size of foveal area stimulated reduces *LDR* effect on $F_{max}$ (25).

4. Decreasing the dark interval in *LDR* has the same effect as reducing visual area with uniform fields (9).

### M. INTENSITY; BRIGHTNESS

1. By proper manipulation of *I* and *LDR*, human *CFF* threshold can be obtained between 1½ and 63 *cps*. Theoretically by proper manipulation of *LDR*, *F* could be increased beyond 63. With *F* and *LDR* fixed one and only one $I_m$ occurs. *CFF* depends on the suppression of intensity discrimination (19).

2. At a constant temperature, with increasing $I$ and $F$, an $I$ is reached beyond which further $I$ increases have no effect, but allowing a longer dark interval for decay of after-image will permit higher $F$'s to be obtained (5).

3. Recognition of flicker requires the establishment of a constant percentage of difference between the total effect of a flash and the after-image (persistence of vision) effect (6).

4. In a single flash, recognition of the effect is one of physical intensity determination rather than of subjective brightness or duration. Recognition of flicker is a form of intensity discrimination, however critical flash intensity cannot be determined by brightness alone. Brightness is one kind of sensory effect while $F_{max}$ is another. Subjective brightness and the factor of the "size of the population" of units of response are not necessarily concurrent nor have they been found to be so all of the time (17, 26).

5. The difference in subjective brightness at *CFF* threshold can be seen by separating the binocular image on arrival at the binocular fusion point; then both the $R$ and $L$ image will be seen to be flickering and to be of unequal apparent brightness (17).

6. At the lower range of $F$ (less than 8 *cps*) and of $\log I$ with a small test patch subjective differentiation of flicker from fusion can be made without the consciousness of any brightness, pattern, or color (23, 26).

7. Since the effect of a surround is not the same at all levels of $I$, one should avoid using surrounds (5).

8. The downward inflection of the upper end of the $F$-$\log I$ curve with larger areas and greater $I$ is primarily due to rod interaction (19).

### N. FREQUENCY: TIME-ENERGY RELATIONS

1. In *CFF* threshold measurement the magnitude of the "sensory effect" must be taken as directly proportional to $F$, but this measures the discrimination of the effect of the flash intensity from the effect of the flash after-image (17).

2. Is the dark time ratio proportional to $\log F$? "This could easily be the result if the dynamic equilibrium in percentage of non-refractory units is the outcome of opposed unlike processes respectively due to excitation and recovery of units in which the intrinsic excitability fluctuates" (15).

3. At any given moment, with a slightly larger number of excited effectors active, the duration of the flash at constant $I$ need not be quite so long to produce the same magnitude of effect and therefore the intensity decay curve (dark time interval) may be decreased (19).

4. Differences in $F_m$-$\log I$ curves as related to $\sigma_F$ correspond to fluctuations in heart beat, body temperature, or metabolism and the like (6, 26).

## O. Talbot's Law

1. The interrelations of flash intensity, fused intensity (Talbot), and subjective brightness are resolved by the relation of the striation of the field to *LDR*. The basic and most important contributing factor is the physical intensity of the flash, rather than its duration or subjective brightness, at least for *LDR*'s over .25 (26).

2. Talbot's law should not be applied in *LDR* studies since it obscures the *L* (flash) and *D* (dark interval) effect as well as giving rise to unsymmetrical *F-log I* curves (19).

3. If Talbot's law held in the interrelation of *F, I,* and *LDR,* then equating *I* for each *F* should produce closer contours as *LDR* is increased; instead they are spread apart (9).

## P. Light-Dark Ratio

1. Decreasing *LDR* with uniform field leads to increase in $F_{max}$, decrease in tau prime, while the slope remains constant. The entire *F-log I* curve is moved to the left on the abscissa and upward on the ordinate as *LDR* decreases (19).

2. Generally speaking the *LDR* effects are dependent on the duration of the dark interval. An increase in $t_D$ permits each flash to act on a larger number of non-refractory sensory units because the chances of recovery are greater with lengthened dark intervals (15).

3. $F_{max}$ of the cone *F-log I* curve is directly proportional to $t_D$ while slope is inversely proportional to $t_D$. The rod effector unit population is less responsive to *LDR* changes than is the cone (25).

4. Test patch striation causes the slope to vary with changes in *LDR* (26).

## Q. Temperature

1. Tau prime is a function of temperature; $F_{max}$ and slope are not (15).

2. Increasing the temperature of the organism increases the metabolic reaction velocities responsible for the capacity being tested. Excitability of visual receptors is measured by $1/I$. The total number of potential response elements is independent of temperature, but the ease of their excitability in terms of $I$ is dependent on temperature. Hence $F_{max}$ and slope are unaffected by temperature but if *LDR* is altered the slope changes since the recovery time available is altered (15).

### R. GENETICS

1. In hybrid sunfish there is shown a simple genetic dominance for the determination of the slope parameter and genetic blending for $F_{max}$ and tau prime. This is evidence that $F_{max}$, slope, and tau prime are the correct parameters of the *F-log I* curve. It is also evidence of the statistical determination rather than the photochemical determination of the *F-log I* curve (6, 13).

### S. BINOCULAR V. MONOCULAR EFFECTS

1. The complete *F-log I* curves for *R, L,* and *B* (WJC) indicated that *B* differs from *R* and *L* in a slight but significant and systematic fashion. In the "cone" portion $I_m$ is almost exactly intermediate between *R* and *L*. For EW the *B, R,* and *L* had different shapes from WJC but *B* was usually intermediate (17).

2. The ratio of monocular (most sensitive eye) to *B* is 1:1.4 with respect to *F* in *cps* (17).

3. The *R, L,* and *B* curves are independent of hue, *LDR,* and subjective brightness level (17).

4. In comparing Crozier's *R, L,* and *B* curves (uncomplicated by pupillary reflex changes in accommodation of adaptation) the *L* and *R* subjective brightnesses are not equal at *CFF* threshold, *R* having a lower threshold. Subjective brightness for *B* is above that of *R* and *L* yet *R CFF* is lower than *L* with *B* intermediate. Hence relation between *F* and *log I* is not determined by subjective brightness alone; this is also true for hue and *LDR* (17).

5. For *B,* the *CFF* of the number of potentially responsive units is doubled as the variation indices prove and the fusion brightness is somewhat increased, but this does not materially increase $F_{max}$. In other words the potential effectiveness of each neural element is doubled but the total number of units of effect is nearly the same (17).

### T. DIVISION OF FIELD; STRIATION: PECTEN EFFECT

1. Doubling the light-dark perimeter (edge) in a striated test patch increased the slope in the ratio of 1:1.4 and upset the regular relation of $t_D$ to slope (26).

2. The important factor in a striated test patch is not the total area illuminated but the unilluminated zones between patches, that is the dark areas (26).

3. The striation of the test patch causes the "rod" portion of the

*F-log I* curve to displace to the left on the abscissa with respect to the complete slope (26).

4. In a subdivided or striated test patch, when *LDR* exceeds .50 the subjective end point of flicker is predominantly along the edges of the area, which means that the number of available cones is increased with consequent reduction in slope (24).

5. The effect of a divided or striated field is to shift the subjective appearance of the field toward lower *I* effects. The outlines are clear down to $F = 18$ (*LDR* .10) and $F = 1$ (*LDR* .90). At lower *I*'s the field has a granular or frosted appearance (24).

6. With an obliquely barred field and rotating stripes the *LDR* relations were about the same as with an open field for .10, .25, and .50; but for .75 and .90 $F_{max}$ and tau prime were sharply changed (26).

## U. HUE

1. Hue, at a given *LDR,* does not vary regularly any one of the three parameters. The "cone" slope is a function of hue only insofar as it is associated with the rate of change of $F_{max}$ with *LDR* (20).

2. By adjusting *LDR* to bring the various hues to the same $F_{max}$ it was possible to compute tau prime for each color band. In the case of Crozier's *L* eye a computed $F_{max}$ of 58.5 was achieved by a computed *LDR* of .91 for Violet, .695 for Blue, .878 for Green, .37 for Red, and .25 for White. The respective tau primes were then $V = -1.31$, $Bl = -3.66$, $G = -2.72$, $R = -2.63$, and $W = -1.20$. On this basis thermopile determinations of flash energy gave $V = -2.76$, $Bl = -2.39$, $G = -2.94$, $R = -2.23$, and $W = -2.60$, indicating that each color band differs in the energy requirement necessary for *CFF* at tau prime (20).

3. $F_{max}$ is higher and tau prime is lower for a given *LDR* for *V, Bl,* or *G* than the white light from which the colors are filtered (20).

4. It cannot be asserted that red or blue acts on 1/3 or 1/4 of cone units if white acts on all of them (20).

## V. SENSORY CORRELATES OR EFFECTS

1. Subjective appearance of the test patch at *CFF* threshold at:
    (*a*) Lowest flash *I*,—smooth bluish grey.
    (*b*) Lower flash *I*,—blue-grey speckled with yellowish flecks.
    (*c*) Medium low flash *I*,—granular grey with many yellow spots.
    (*d*) Medium high flash *I*,—frosted yellowish.
    (*e*) High *I*,—smooth surface, sharp outline of edges.
    (*f*) High *I* (if colored),—smooth surface of color.

For *LDR* .10 step (*e*) is about 35 *cps; LDR* .25 about 21 *cps, LDR* .50 about 20 *cps, LDR* .75 about 19 *cps* and *LDR* .90 about 16 *cps* (19).

2. Subjectively the *CFF* of a field, subdivided or not, is seen to flicker or fuse all at once but the effect is sharper at the boundaries or edges. When *LDR* is greater than .50 the subjective end point of flicker is predominantly along the edges (24, 26).

3. The change in subjective appearance of the field at *CFF* threshold cannot possibly be used to differentiate the entrance of cone elements or absence of rod units (23).

4. The parameters of the *F-log I* curve are independent of subjective color. There are no discontinuities or changes when hue becomes apparent as *I* increases (23).

5. With a striated test field *F-log I* contours for subjective phi and gamma as well as flicker-fusion were obtained. Gamma is the appearance of pulsing, widening, and contraction of the stripes at a rate lower than of the flashing. Phi is a pulsing side-to-side movement of the entire field (26).

## W. Critical Comments

1. Crozier and Wolf stated that theirs was an "intensity-discrimination theory of the marginal recognition of flicker." If the determination of the *CFF* threshold is an intensity discrimination function, namely the judgment of equality or non-equality of two patterns as to apparent brightness, this can be true only in very special and limited cases. All of their moving stripe experiments with animals would seem to indicate that the recognition or non-recognition of movement was the adequate stimulus. In their discriminometer experiments they reported that the flicker-fusion point is a smooth uniform field. In $F_m$ experiments the apparent brightness is governed by Talbot's law; in the $I_m$ experiments the shift in *I* surely was frequently below a single unit of ΔI. Their basic assumption gained them little or nothing.

2. The fact that the *F-log I* curve has the shape of a probability integral enabled them to apply probability statistics to their data throughout its range but there is no evidence that the application of statistical mechanics to any but the measurement of single points gained them anything. In one place where it did permit an empirical verification, namely equating colors for their basic physical intensity effect, they did not carry through an empirical check.

3. The parameters of the *F-log I* curve, $F_{max}$, tau prime, and slope are logical *if* it is a probability integral. The height of the curve, its slope,

and its inflection point would seem to be sufficient to describe the curve. However each succeeding experiment showed that these three parameters had neither constant nor logical, structural, or functional correlates. Indeed these parameters led them into numerous paradoxes and absurdities. They never stated their method of determination of $F_{max}$, hence one cannot check from their published data the accuracy of their computed parameters.

4. Their argument and demonstration of the limitations of Hecht's photochemical theory of retinal function seems sound. Hecht's theory was a limited one. It had value in that it led on to the knowledge concerning the chemistry of visual purple and vitamin *A*. Hecht generalized too far from his theory.

5. Their use of the term "after-image" was unfortunate since they were referring to the persistence of visual sensation in the dark interval ($t_D$). After-image has for many years been the customary designation of a recurrent image appearing one or more seconds after the physical stimulus has ceased.

6. Crozier and Wolf did a larger number of precise, well-designed, and controlled experiments basic to the *CFF* threshold phenomenon than anyone else who has worked in this field. They made several original discoveries. Their conceptual framework of explanation was not one which enabled them to present their findings in an integrated form.

## REFERENCES

1. CROZIER, W. J. On the variability of critical illumination for flicker fusion and intensity discrimination. *J. Gen. Physiol.*, 1936, **19**, 503-523.
2. CROZIER, W. J., WOLF, E., & ZERRAHN-WOLF, G. On critical frequency and critical illumination for response to flickered light. *J. Gen. Physiol.*, 1936, **20**, 211-228.
3. ———. Critical illumination and critical frequency for response to flickered light, in dragonfly larvae. *J. Gen. Physiol.*, 1937, **20**, 363-392.
4. ———. Temperature and critical illumination for reaction to flickering light: I. *J. Gen. Physiol.*, 1937, **20**, 393-410.
5. ———. Temperature and critical illumination for reaction to flickering light: II. *J. Gen. Physiol.*, 1937, **20**, 411-431.
6. ———. Critical illumination and flicker frequency in related fishes. *J. Gen. Physiol.*, 1937, **21**, 17-56.
7. ———. Intensity and critical frequency for visual flicker. *J. Gen. Physiol.*, 1937, **21**, 203-221.
8. ———. Critical illumination for response to flickered light with dragonfly larvae (Anax) in relation to area of eye. *J. Gen. Physiol.*, 1938, **21**, 223-248.
9. ———. Critical illumination and flicker frequency, as a function of flash duration for the sunfish. *J. Gen. Physiol.*, 1938, **21**, 313-334.

10. ———. Critical intensity and flash duration for response to flicker; with Anax larvae. *J. Gen. Physiol.*, 1938, **21**, 463-474.

11. ———. The flicker response function for the turtle Pseudemys. *J. Gen. Physiol.*, 1939, **22**, 311-340.

12. Crozier, W. J., & Holway, A. H. Theory and measurement of visual mechanisms: I. *J. Gen. Physiol.*, 1939, **22**, 341-364.

13. Crozier, W. J., & Wolf, E. The flicker response contours for genetically related fishes: II. *J. Gen. Physiol.*, 1939, **22**, 463-486.

13a. ———. The flicker response contour for the gecko (rod retina). *J. Gen. Physiol.*, 1939, **22**, 555-566.

14. Crozier, W. J., & Holway, A. H. Theory and measurement of visual mechanisms: III. *J. Gen. Physiol.*, 1939, **23**, 101-141.

15. Crozier, W. J., & Wolf, E. Temperature and critical illumination for reaction to flickering light: VI. *J. Gen. Physiol.*, 1940, **23**, 531-551.

16. ———. Reaction to visual flicker in the Newt Triturus. *J. Gen. Physiol.*, 1940, **23**, 667-676.

17. ———. Theory and measurement of visual mechanisms: IV. *J. Gen. Physiol.*, 1941, **24**, 505-534.

18. ———. The simplex flicker threshold contour for the Zebra Finch. *J. Gen. Physiol.*, 1941, **24**, 625-633.

19. ———. Theory and measurement of visual mechanisms: V. *J. Gen. Physiol.*, 1941, **24**, 635-654.

20. ———. Theory and measurement of visual mechanisms: VI. *J. Gen. Physiol.*, 1941, **25**, 89-110.

21. ———. Theory and measurement of visual mechanisms: VIII. *J. Gen. Physiol.*, 1942, **25**, 369-379.

22. ———. The wavelength sensitivity for the Zebra Finch. *J. Gen. Physiol.*, 1942, **25**, 381-390.

23. ———. Theory and measurement of visual mechanisms: IX. *J. Gen. Physiol.*, 1943, **27**, 119-138.

24. ———. Theory and measurement of visual mechanisms: X. *J. Gen. Physiol.*, 1944, **27**, 287-313.

25. ———. Flicker response contours for the sparrow, and the theory of the avian pecten. *J. Gen. Physiol.*, 1944, **27**, 315-324.

26. ———. Theory and measurement of visual mechanisms: XI. *J. Gen. Physiol.*, 1944, **27**, 401-432.

27. ———. Theory and measurement of visual mechanisms: XII. *J. Gen. Physiol.*, 1944, **27**, 513-528.

28. Wolf, E., & Zerrahn-Wolf, G. Reactions of limulus to illuminated fields of different area and flicker frequency. *J. Gen. Physiol.*, 1937, **20**, 767-776.

*Psychiatric Institute*
*722 West 168th Street*
*New York City 32*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 19-25.

# VARIABILITY AS A VARIABLE*

*Department of Psychology, Pennsylvania State College*

---

WILLIAM M. LEPLEY

---

## A. INTRODUCTION

Biological and social scientists are interested in predicting the behavior of organisms, other organisms. Organisms are *interested* in being unpredictable. This latter is the central thesis of this writing. Restated, and with the non-technical use of the term *interested* avoided, the proposition becomes: It is biologically important that the behavior of organisms be unpredictable; behavior variability has survival value; behavior variability is an aspect of, and a contribution to, adaptability. This proposition seems to be especially applicable in situations wherein organisms stand in the relation of antagonists; the hunter and the hunted; the oppressor and the oppressed; paired debaters, painted fencers, paired chess players. Variability of behavior is a mode of attack and a mode of defense. Furthermore the concept seems to apply to situations wherein the organism is pitted against the non-biological, non-social environment as antagonist; the cat in a puzzle-box, man in an adverse climate, a student with a problem to solve.[1]

This variety of variability has been pointed out and much discussed by students of trial-and-error learning. The stereotyped response is likely to be ineffective. Variability of behavior is adaptive in that a *correct* response is ultimately hit upon. From this point onward, many theorists have proceeded to describe and interpret the learning process. As early examples, both Bain and Spencer described the formation of habits as a process of selection from a varied repertoire of movements. It is not the purpose of this writing to develop learning theory, but to examine and interpret variability itself; to point the finger of research at variability rather than consistency; at the standard deviation rather than the mean; range rather than mode; to propose and emphasize that behavior variability may be thought of, and dealt with, as lawful rather than capricious. More specifically it sets forth a simple theoretical framework which is intended to call attention to intra-individual variability as an interindividual variable.

---

*Received in the Editorial Office on July 2, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The writer sees no objection to extending the notion to antagonistic social structures; and to the ecology of societies.

One useful variety of hypothesis is the type that proposes to bring together, in one larger family of phenomena, the phenomena of two or more families theretofore considered and interpreted separately. What is new in this type of formulation is the simpler theoretical structure and the larger family. If clear-cut, such hypotheses usually generate questions and problems to be answered and solved by a more systematic examination of the phenomena of each of the families included. The hypothesis proposed in the foregoing and restated below is judged to be of this type. It is proposed that: Behavior variability has survival value; that it is an aspect of, and a contribution to, adaptability. It is suggested that behavior variability may be a fundamental or near-fundamental property of protoplasm.[2]

## B. THE GENERATION OF BEHAVIOR VARIABILITY

All varieties of writings about the behavior of organisms are replete with suggestions and explicit statements that acts of behavior generate barriers to their reoccurrence; that organisms tend to be non-repetitive; that they avoid and resist repetitiousness; but it has been only relatively recently that the general principle has been formulated and made explicit. The writer has borrowed the principle of Reactive Inhibition from Hull (6) and from his intellectual ancestors; Pavlov (8), Thorndike (9), Luh (7), Bills (1), Dodge (2), and Dunlap (3), to mention only a few. As used herein, the principle is proposed as that which underlies the generation of behavior variability. The following statements are quoted from Hilgard (5): "The evocation of any reaction generates reactive inhibition. Reactive inhibition is spontaneously dissipated in time." The preceding quotation is a modification of Hull's (6) Postulate No. 8, which is quoted, in part: "Whenever a reaction (R) is evoked in an organism there is created as a result a primary negative drive (D); (*a*) this has an innate capacity ($I_R$) to inhibit the reaction potentiality ($_SE_R$) to that response; (*b*) the amount of net inhibition ($I_R$) generated by a sequence of reaction evocations is a simple linear increasing function of the number of evocations (n); . . . (*d*) reactive inhibition ($I_R$) spontaneously dissipates as a simple negative growth function of time (t ' ' ')."

The writer proposes to substitute the following modifications and paraphrasing for the foregoing.[3] When any reaction occurs, the probability of its later occurrence is reduced. The probability of its later occurrence increases with the continuing succession of following dissimilar reactions. It is proposed that the probability of later occurrence would be reduced progressively with continued repetition; and that the degree of reduction would be proportional to the degree of repetitiousness or, in other words, and inversely, to the number and spacing of interpolated dissimilar reactions. It is further proposed that this reduction in the probability of occurrence for a particular reaction would enhance the probability of occurrence for other reactions in the repertoire (or even novel reactions).

---

[2] It should be admitted that extremes of behavior variability may be maladaptive. This seems to be especially true in all non-hostile social behavior. The worker whose arrival and performance at work is unpredictable is a likely candidate for discharge. Civilized living requires compromise and norming along this continuum, as well as for most variables.

From these laboriously expressed propositions may be extracted the simple statements that behaving generates behavior variability; the behavior of organisms tends to be variable.[4]

## C. Some Implications

This and the following sections are written to propose and discuss a limited number of tests of the hypothesis under consideration. A statement of one form of the general case follows. If it is true that behavior variability is an aspect of, and a contributor to, adaptability, then it follows that deficiencies in, or impairments of, adaptability might be reflected in behavior as reduced or restricted variability. The following paragraphs deal with a few of the possible special cases.

### 1. *Amentia*

This old term refers to congenital deficiency in intelligence; and the concept of adaptability is a part of all modern definitions of intelligence. In this first special case then, it should be possible to show that individuals deficient in intelligence exhibit restricted variability in their behavior. As evidence the writer calls attention to the long-accepted finding that the vocabulary test score is one of the better single indices of intelligence. The large vocabulary is interpreted here as a richly variable behavior repertoire.[5]

### 2. *Age*

It is not meaningful to speak of age as such. The reference here is to organic changes that take place in time. It is believed to be generally

---

[3] It will be noted that the variable, Work (W), has been omitted from this paraphrase.

[4] It is judged possible that, at least in part, this principle may account for what Hull (6) has called "behavioral oscillation."

[5] The writer has underway a project designed to test the hypothesis that the large synonym vocabulary is an especially valid indication of high intelligence.

11·8·70
J 112

agreed that adaptability changes throughout the life span; first increasing from infancy to maturity and later declining in old age. It should be possible to show that behavior variability covaries with these trends. The writer submits as evidence the casual observations of the repetitious infant and the highly predictable dotard.

### 3. *Toxic States*

The writer is not prepared to estimate the adaptability impairment by toxins in general. However, he is somewhat familiar with the toxic state brought about by the ingestion of ethyl alcohol, and he wishes to invite attention to the persistent, perseverating, obsessive, indistractable inebriate.

### 4. *Disturbed and Excited States*

This category is in need of subdivision and the subdivisions are not easily accomplished by the present writer; this because of the lack of clear definition in matters psychological having to do with emotional behavior, motivation and similar. One somewhat defensible subclassification is that which separates transient or acute states from the chronic. Such is proposed and used here.

### 5. *Chronic*

A further subdivision frequently used is the one which sets the neurotic and the psychotic apart.

*a. Psychotic.* If psychosis is thought of as an impairment of adaptability then, according to the propositions being examined, one might expect restricted behavior variability in the psychotic. The compulsions, obsessions, and unshakeable systematic delusions of the psychotic are too familiar for more than mention.

*b. Neurotic.* Common, recurring terms used in the description of the neurotic life-style are: rigid, perseverative, compulsive, over-conforming, etc.

### 6. *Transient or Acute*

Proposed for this category are various emotional-motivational crises.

*a. Angry states.* Outstandingly relevant here is the regressive, abortive stereotyped attack as contrasted with the planned, varied and, in social relations, deceptive campaign of behavior, of the relatively undisturbed organism.

*b. Fearful states.* The general behavior of organisms in transient, fearful states is so similar to behavior in angry states as to require no separate description. An organism critically afraid is relatively more predictable.

*c. Other, heightened motivational states.* The writer ventures the generalization that any sufficiently radical disequilibrated state results in, or is accompanied by, restriction and stereotyping in behavior, both overt and covert.

*d.* Of especial current interest to the writer is the progressive predictability (stereotyping) that appears to characterize the behavior of organisms as they approach closer and closer to goals.[6] Fajans (4) has discussed this phenomenon in her observations of the behavior of young children. A familiar, personal example to the writer is noted in his generally inadequate end-game in chess. As he nears the goal he suffers a perceptual and conceptual rigidity that is likely to leave him highly vulnerable. Another familiar example is seen in the increasing compulsiveness of the graduate student as he approaches nearer and nearer to his academic goal; the coveted Ph.D. The writer proposes these examples as essentially similar to the running behavior of a rat in a maze. He has some preliminary data which appear to show that rats run the near-goal segments with a higher reliability. Outside of the protective laboratory such predictability might make the animal easy prey to his enemies.

## D. FURTHER IMPLICATIONS

A second formulation or general case is proposed as follows: If it is true that behavior variability is an aspect of, and a contributor to, adaptability, then it follows that adaptable organisms may tend to resist, and to escape from, restrictions externally imposed upon the behavior repertoire. The voluminous technical literature on boredom and monotony testify to this point; and throughout this literature one finds the suggestion that repugnance to repetitiousness covaries with adaptability or intelligence. Wherever apparent exceptions occur, at least two simple interpretations seem reasonable. The intelligent organism may, in some sense, subjectively impose or superimpose variability upon repetitiousness. There is nothing original in this interpretation. Second; intelligent organisms suffering disturbed states or heightened motivational states, etc. (see foregoing section) may exhibit enhanced tolerance with regard to restrictive circumstances. The writer doubts that a starving man would complain about the repetitive boarding-house menu;[7] that a frantic fugitive would be bored while passing through monotonous scenes; that a man shouting in anger would display a rich synonym vocabulary, or be unduly annoyed by the triteness of phrasing in the responses of his antagonist. Disturbed organisms are comforted and relieved in refuges among familiar things and permitting routinized living.

---

[6] Certain critics of this paper have insisted that the principles set forth can apply only to goal-oriented behavior. The writer at present, prefers to maintain the postulate that all behavior is goal-oriented; just as he would insist that all behavior (or all activity and change to include the inorganic) has dynamic antecedents and accompaniments. "Behavior without a goal" is to him an unacceptable phrase; and he would suggest that the search for the goals of so-called automatisms, etc., to be continued vigorously.

It would be a simple exercise to reformulate the principles foregoing in this paper and produce a fairly coherent discussion of curiosity behavior.[8] Anabolic processes generate energy. Organisms have energy to spend. Under certain conditions, the reactive inhibition principle insures that they not expend it all in one way and place; hence curiosity behavior; doing this then trying that; going thither then yon; perhaps only apparently "seeking," but with the same adaptive outcomes. Consistent with earlier discussions, one might expect curiosity behavior to be depressed among the stupid; among the disturbed and anxious; for mankind, among those who live insecurely in an atmosphere of hostility. The cat killed by curiosity is judged to have lived and died under a dictatorship. This biological argument is extendable for the clear support of the case for human freedoms.

It is judged unnecessary to elaborate further these paragraphs on the tendencies of adaptive organisms to resist stereotypy in behaving and perceiving; to escape from it and to modify it in the direction of variation. These matters are so familiar to all. The writer would not further commit the literary error of repetitiousness about repetitiousness.

## E. Summary

There is nothing in this paper other than the amplification, exemplification, and repetition of the hypothesis that: Behavior variability has and has had survival value in the evolutionary sense; that it is an aspect of, and a contribution to, the adaptability of organisms; that it may be thought of, and dealt with, as a lawful variable.

## References

1. Bills, A. G. Blocking: a new principle of mental fatigue. *Amer. J. Psychol.*, 1931, **43**, 230-245.
2. Dodge, R. Note on Professor Thorndike's experiment. *Psychol. Rev.*, 1927, **34**, 237-240.

---

[7] The writer is engaged in a series of experiments intended to test the hypothesis that certain organisms can be induced to increase their food intake by varying only the visual aspects of the food.

[8] It should be admitted, of course, that, above the primitive level, curiosity behavior is modificable in accordance with learning principles.

3. Dunlap, K. Habits, Their Making and Unmaking. New York: Liveright, 1932.
4. Fajans, S. Die Bedeutung der Entfernung für die Stärke eines Aufforderungscharakters beim Säugling und Kleinkind. *Psychol. Forsch.*, 1933, **17**, 215-267.
5. Hilgard, E. R. Theories of Learning. New York: Appleton-Century, 1948.
6. Hull, C. L. Principles of Behavior. New York: Appleton-Century, 1943.
7. Luh, C. W. The avoidance of repetitive responses. *Psychol. Rev.*, 1931, **38**, 276-279.
8. Pavlov, I. P. Conditioned Reflexes. (Trans. by G. V. Anrep.) Oxford: Oxford Univ. Press, 1927.
9. Thorndike, E. L. The refractory period in associated processes. *Psychol. Rev.*, 1927, **34**, 234-236.

*Department of Psychology*
*Pennsylvania State College*
*State College, Pennsylvania*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 27-30.

# PURE GUIDANCE AND HANDLING AS COMPONENTS OF THE MAIER TECHNIQUE FOR BREAKING ABNORMAL FIXATIONS*

*Department of Psychology, University of New Hampshire*

G. M. HASLERUD, L. BRADBARD, AND R. P. JOHNSTON[1]

## A. THE PROBLEM

Maier has postulated that fixated behavior resulting from frustration is qualitatively different from position habits which arise during the course of trial and error learning. Manual guidance was found by him to be an effective cure of (3) and also prophylactic (5) for such fixations. That Maier may have overvalued the pointing out of alternatives to the animal has been suggested by Haslerud (1) who found that while mechanical guidance is just as effective as trial and error for original anticipative learning in a free maze, it is not at all useful for changing anticipations. If the fixations on the jumping apparatus have a similar law, then by separating Maier manual guidance into pure (mechanical) guidance and handling, a test can be made of how necessary the guidance is. Only if the mechanical guidance group is best or if the groups with manual and mechanical guidance are equally good and the handling (trial and error) group poorer, can the value of guidance be maintained. If all are undifferentiated, then probably neither the pointing out of the correct alternative nor simple handling is responsible for the results.

## B. PROCEDURE

A Lashley jumping apparatus was modified by a small slot in the nosepiece between the windows and a notch at each of the front corners of the square jumping platform. A 10 x 11 in. transparent "Lucite" barrier slung between the nosepiece and the platform allowed the mechanical guidance group to see but prevented them from jumping to the negative card. The barrier was removed for the other two groups. To link in a straight-line relationship the discrimination situation and the full reward, the individual feeding cages were placed directly behind the feeding platform of the apparatus. The rats came with 23-hour hunger to the daily 10 massed trials.

---

*Received in the Editorial Office on July 2, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Data is a part of that in M.A. theses by the junior authors. This report is a collaboration with final responsibility the senior author's.

The procedure involved the setting up of a fixation before the discrimination problem was introduced, as in Maier and Feldman (4). During the preliminary 10 days the rats learned to jump 8½ in. to knock down gray cards at the windows, manual guidance equalizing jumps to each side. Then in Part I for 16 days the regular Maier stimulus cards, black circle on white and the reverse, were placed in the windows randomly and one or the other window locked randomly. With jumps to each card reinforced with food 25 per cent of the time and punished by falling 25 per cent, the animals fixated a position response which tended to give them 50 per cent success. Finally in Part II for a maximum of 20 days the card with black circle and white background became consistently positive, and the other card was always locked. The abandoning of the fixation and the perfecting of the discrimination to a criterion of 29 out of 30 alternate free trials were noted especially. In Parts I and II the animals were put on the jumping platform backward to obviate pointing.

Before beginning Part II the rats were separated into the three experimental groups equated on the basis of their emotionality in Part I as judged by proneness to abortive jumps, defecations, and urinations. The guidance for the Maier and mechanical guidance groups and the special handling, i.e., emotional soothing for 30 sec., for the third group were given on the odd trials one day and the even trials the next. In order to control the amount of handling, all the animals on the non-guided trials and the Maier and mechanical guidance groups on the guided trials were transported from home cages and shifted between trials by a metal grain scoop. To test whether the retention of discrimination was different for the three groups, each rat was given 30 unguided free trials after two weeks of rest following criterion on Part II.

The experiment began with 29 male white rats about five months old which were naïve for discrimination problems and jumping but were accustomed to handling from an earlier maze experiment. Not all the animals could be used, just as Maier and others have found. Six had to be eliminated because of abortive jumping on 50 per cent or more of the days in Part II; only those animals jumping directly at the cards can be studied for development of discrimination. An additional four animals were too resistant to jumping. If a rat required tail taps (5) on more than one trial out of the 10 in a day for 50 per cent or more of the days of both Parts I and II, it was considered possibly motivated by "anxiety" or escape from punishment rather than by the food behind the positive card. Seven rats were left in each of the manual and mechanical guidance groups and five in the handling group. All of these met the criterion of learning.

## C. RESULTS

Inspection of Table 1 reveals the three groups not very different. This is verified by analysis of variance. The $F$'s are only .45 and 1.04, obviously non-significant for 2 and 16 df, for trials required for first break from the stereotyped jumping which had developed in Part I and the additional trials to the beginning of the criterion trials respectively. Similarly the retention was equally high for all groups, showing independence of extraneous aids.

TABLE 1
MEASURES OF CENTRAL TENDENCY (MEDIAN IN PARENTHESES) FOR DROPPING STEREOTYPY AND FOR LEARNING AND RETENTION OF DISCRIMINATION

| Groups | Maier manual guidance | Mechanical guidance | Handling |
|---|---|---|---|
| *N* | 7 | 7 | 5 |
| Trials for breaking fixation | 22.4 (17) | 34.1 (30) | 29.8 (28) |
| Additional trials to learn to beginning of criterion | 76.3 (53) | 47.7 (42) | 46.8 (55) |
| Correct responses on 30 free retention trials | 28.3 (29) | 28.9 (28.9) | 30 (30) |

The three groups had been equated for "emotionality" at the end of Part I. At the end of Part II after the different kinds of guidance and handling, there was again no differentiation between the groups but rather a large decrease in the urination and defecation scores for all groups.

## D. DISCUSSION

Our results for the manual guidance group agree well with those in Maier and Klee (3). Their 15 rats with guided alternate trials showed a mean of 32.7 trials to abandon position stereotype and 69.2 for additional trials before reaching criterion run as compared with our 22.4 and 76.3 respectively. With similar techniques and similar results, generalizations can therefore go further than for our small group.

Since no significant difference could be found between our three groups for either trials to break fixation or additional trials to learn, some factor that is neither pure guidance nor simple handling must be common to all three group situations. What might such a factor be?

It is probably a collective "conditions favorable for learning," as has also recently been emphasized by Wilcoxon (6). By selection of good jumpers and by direct linking of positive card to food, the motivation is toward food rather than away from punishment. The change from insoluble Part I to soluble Part II is basically raising the reinforcement of the positive card to 50 per cent from the 25 per cent of Part I and reducing the negative card to zero. The guidance groups would have in addition on the guided trials 100 per cent reinforcement. However, the studies of partial reinforcement (2) do not show 100 per cent superior to 50 per cent in establishing response strength. Certainly our guidance groups manifested no advantage from the additional reinforcement in either learning or retention. By extrapolation, one might wonder if the guidance feature of certain psychotherapies is nearly as important as more fundamental conditions of learning.

## REFERENCES

1. HASLERUD, G. M. Anticipative transfer of mechanically guided turns. *J. Exp. Psychol.,* 1953, **45**, 431-436.
2. JENKINS, W. C., & STANLEY, J. C. Partial reinforcement: a review and critique. *Psychol. Bull.,* 1950, **47**, 193-234.
3. MAIER, N. R. F., & KLEE, J. B. Studies of abnormal behavior in the rat: XVII. Guidance versus trial and error in the alteration of habits and fixations. *J. of Psychol.,* 1945, **19**, 133-163.
4. MAIER, N. R. F., & FELDMAN, R. S. Studies of abnormal behavior in the rat: XXII. Strength of fixation and duration of frustration. *J. Comp. & Physiol. Psychol.,* 1948, **41**, 348-363.
5. MAIER, N. R. F., & ELLEN, P. Studies of abnormal behavior in the rat: The prophylactic effects of "guidance" in reducing rigid behavior. *J. Abn. & Soc. Psychol.,* 1952, **47**, 109-116.
6. WILCOXON, H. C. "Abnormal fixation" and learning. *J. Exp. Psychol.,* 1952, **44**, 324-334.

*Department of Psychology*
*University of New Hampshire*
*Durham, New Hampshire*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 31-38.

# SPREAD AND INTENSITY OF VOCATIONAL INTERESTS AND EVALUATIVE ATTITUDES IN FIRST-YEAR NEGRO MEDICAL STUDENTS*

*Department of Psychology, Howard University*

ALBERT T. MILAM AND F. C. SUMNER

It is proposed in the present study to ascertain in a group of first year Negro medical students *(a)* the intensity of physician interest; *(b)* the spread of vocational interests and evaluative attitudes of significant intensity; (*c*) the relationship of spread and intensity of such vocational interests and evaluative attitudes to the scholarship level in the medical studies of the first year.

## A. METHOD

The Strong Vocational Interest Blank for Men and the Allport-Vernon Study of Values were administered through February and March, 1950, to the 1949-50 First Year medical class of the Howard University Medical School ($N=74$ less 5 Negro women students and 3 white male students). The Strong was machine-scored for the 39 vocations and all Strong scores are here reported in terms of standard scores.

Medical school grades for the entire first year were obtained and translated into individual honor point averages for the year's work with the following results for the 66 members of the class: 2.67 through 2.00 was the obtained *B* range in honor point averages ($N=8$); 1.92 through 1.00, the *C* range ($N=38$); .917 through 0, the *D* range, in which two subjects who were the only ones making below zero were included for convenience of statistical treatment, bringing to 20 the *N* of the *D* group.

## B. RESULTS

In Table 1 are shown for the first year medical students: *(a)* the mean Strong scores for the group as a whole in each of the 39 vocational interests; *(b)* the per cent of the class as a whole making significantly high Strong scores *(B+* and *A)* in each of the 39 vocational interests; *(c)* the mean Strong scores of the *B, C, D* grade groups respectively in each of the 39 vocational interests; *(d)* the number and per cent of each grade group *(B, C, D,* respectively) making significantly high Strong scores *(B+* and *A)* in each of the 39 vocational interests.

---

*Received in the Editorial Office on July 7, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

According to the Strong Vocational Interests norms, standard scores 39.5 through 44.4 equal *B*+ or high positive vocational interest while standard scores 44.5 through 75.4 equal *A* or very high positive vocational interest intensity.

In Table 1 the salient points appear to be the following:

1. In terms of mean standard scores (Strong) of the class as a whole for each of the 39 vocations, the interests, while scattering over the entire gamut of vocations to some degree of intensity, are highly significant only in two vocations, namely, Osteopath (43.8) and Mathematics-Physical Science Teacher (42.7), and in these only at *B*+ intensity.

2. While the mean interest score of the class as a whole for Physician is 39.3 which is slightly below the threshold of highly significant intensity, it is, however, conspicuously higher than the mean score reported by Hartshorn for Negro medical practitioners.

3. For the class as a whole the per cent making highly significant individual scores (*A* and *B*+) varies from a high of 71 per cent (Osteopath) to a low of 00 per cent (Banker), with 53 per cent (Physician) and the correlation between class averages for the 39 vocational interests and the respective percentages of significant individual scores being .89.

4. While 27.3 per cent of the First Year medical students of the present study attained Strong scores of *A* intensity for Physician, 71 per cent of white practicing physicians and 73 per cent of white senior medical school students attained Strong scores of *A* intensity for Physician, and while 25.8 per cent of the First Year medical students of the present study attained Strong scores of *B*+ intensity for Physician, only 13 per cent of white practicing physicians and only 14 per cent of white senior medical school students attained *B*+ interest scores for Physician according to the report of Strong and Tucker.

5. In terms of school grade level of these first year medical students *B* grade students ($N=8$) registered significantly high mean scores in four vocations: Physician (43.3), Musician (42.6), Mathematics-Physical Science Teacher (42.4), and Osteopath (41.9); *C* grade students ($N= 38$) registered significantly high mean scores in only two vocations: Osteopath (43.7) and Mathematics-Physical Science Teacher (43.0); *D* grade students ($N=20$) registered significantly high mean scores in only two vocations: Osteopath (44.6) and Mathematics-Physical Science Teacher (41.3). All significantly high mean scores at each grade level fall short of *A* intensity save for Osteopath in *D* grade students.

6. Significantly high individual scores (*B*+ and *A*) varied for the *B* grade students from a high of 87.5 per cent (Physician) to a low of 00 per cent for each of the following: Forest Service Man, Certified Accountant,

## TABLE 1

MEAN STRONG SCORES AND THE SPREAD OF SIGNIFICANT INDIVIDUAL SCORES OF 66 MALE FIRST YEAR NEGRO MEDICAL STUDENTS FOR THE GROUP AS A WHOLE AND FOR EACH GRADE LEVEL

| | Entire group $N=66$ | | *B* Grade students $N=8$ | | | | *C* Grade students $N=38$ | | | | *D* Grade students $N=20$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean score | % SS* | Mean score | B+ N | A N | % SS | Mean score | B+ N | A N | % SS | Mean score | B+ N | A N | % SS |
| Artist | 26.3 | 12.1% | 31.6 | 0 | 1 | 12.5% | 25.5 | 2 | 2 | 10.5% | 25.6 | 3 | 0 | 15.0% |
| Psychologist | 31.9 | 19.7 | 37.0 | 2 | 1 | 37.5 | 31.0 | 5 | 2 | 18.4 | 31.6 | 2 | 1 | 15.0 |
| Architect | 28.4 | 16.7 | 33.9 | 2 | 0 | 25.0 | 27.6 | 3 | 2 | 13.2 | 27.8 | 4 | 0 | 20.0 |
| Physician | 39.3 | 53.0 | B+43.3 | 3 | 4 | 87.5 | 38.7 | 10 | 8 | 47.4 | 39.1 | 4 | 6 | 50.0 |
| Osteopath | B+43.8 | 71.2 | B+41.9 | 1 | 3 | 50.0 | B+43.7 | 11 | 17 | 73.7 | A 44.6 | 6 | 9 | 75.0 |
| Dentist | 36.4 | 39.4 | 39.4 | 4 | 1 | 62.5 | 35.8 | 9 | 5 | 36.8 | 36.2 | 2 | 5 | 35.0 |
| Mathematician | 26.0 | 7.6 | 30.1 | 2 | 0 | 25.0 | 25.3 | 2 | 0 | 5.3 | 25.7 | 1 | 0 | 5.0 |
| Physicist | 21.9 | 7.6 | 27.5 | 1 | 1 | 25.0 | 20.5 | 1 | 1 | 5.3 | 22.4 | 1 | 0 | 5.0 |
| Engineer | 30.4 | 22.7 | 33.0 | 2 | 1 | 37.5 | 29.2 | 4 | 3 | 18.4 | 31.8 | 4 | 1 | 25.0 |
| Chemist | 35.1 | 28.8 | 38.6 | 2 | 2 | 50.0 | 33.6 | 5 | 2 | 18.4 | 36.4 | 5 | 3 | 40.0 |
| Production Manager | 31.3 | 19.7 | 31.0 | 1 | 1 | 25.0 | 31.1 | 5 | 2 | 18.4 | 31.8 | 3 | 1 | 20.0 |
| Farmer | 34.5 | 34.8 | 34.1 | 0 | 1 | 12.5 | 33.8 | 8 | 6 | 36.8 | 36.1 | 5 | 3 | 40.0 |
| Aviator | 34.2 | 28.8 | 33.1 | 2 | 1 | 37.5 | 33.7 | 3 | 7 | 26.3 | 35.5 | 4 | 2 | 30.0 |
| Carpenter | 26.4 | 12.1 | 25.1 | 0 | 1 | 12.5 | 25.8 | 3 | 2 | 13.2 | 28.1 | 0 | 2 | 10.0 |
| Printer | 37.3 | 39.4 | 37.6 | 1 | 3 | 50.0 | 37.3 | 3 | 11 | 36.8 | 37.0 | 4 | 4 | 40.0 |
| Mathematics-Physical Science Teacher | B+42.7 | 63.6 | B+42.4 | 3 | 3 | 75.0 | B+43.0 | 4 | 19 | 60.5 | B+41.3 | 7 | 6 | 65.0 |
| Policeman | 35.9 | 34.8 | 32.9 | 1 | 1 | 25.0 | 36.2 | 9 | 7 | 42.1 | 36.7 | 2 | 3 | 25.0 |
| Forest Service Man | 27.3 | 13.6 | 25.9 | 0 | 0 | 0.0 | 27.0 | 5 | 2 | 18.4 | 28.3 | 1 | 0 | 5.0 |
| YMCA Physical Director | 37.7 | 40.9 | 34.9 | 3 | 0 | 37.5 | 38.6 | 6 | 11 | 44.7 | 37.1 | 3 | 4 | 35.0 |
| Personnel Director | 31.9 | 22.7 | 29.5 | 1 | 0 | 12.5 | 32.5 | 5 | 5 | 26.3 | 31.9 | 2 | 3 | 25.0 |

TABLE 1 (contd.)

| | Entire group *N*=66 | | *B* Grade students *N*=8 | | | | *C* Grade students *N*=38 | | | | *D* Grade students *N*=20 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean score | % SS* | Mean score | B+ N | A N | % SS | Mean score | B+ N | A N | % SS | Mean score | B+ N | A N | % SS |
| Public Administrator | 36.9 | 39.4% | 34.3 | 2 | 0 | 25.0% | 37.3 | 9 | 7 | 42.1% | 37.2 | 6 | 2 | 40.0% |
| YMCA Secretary | 28.6 | 16.7 | 25.5 | 0 | 1 | 12.5 | 29.6 | 3 | 4 | 18.4 | 27.9 | 1 | 2 | 15.0 |
| Social Science High School Teacher | 35.6 | 36.4 | 34.6 | 0 | 1 | 12.5 | 36.7 | 9 | 9 | 47.4 | 34.0 | 2 | 3 | 25.0 |
| City School Superintendent | 29.5 | 16.7 | 31.8 | 0 | 1 | 12.5 | 29.7 | 7 | 0 | 18.4 | 27.8 | 3 | 0 | 15.0 |
| Minister | 28.8 | 15.2 | 29.3 | 0 | 1 | 12.5 | 29.6 | 2 | 4 | 15.8 | 27.3 | 1 | 2 | 15.0 |
| Musician | 38.1 | 50.0 | B+42.6 | 2 | 4 | 75.0 | 38.3 | 9 | 11 | 52.6 | 35.9 | 3 | 4 | 35.0 |
| Certified Public Accountant | 29.1 | 4.6 | 30.1 | 0 | 0 | 0.0 | 29.6 | 2 | 1 | 7.9 | 27.8 | 0 | 0 | 0.0 |
| Accountant | 28.9 | 13.6 | 26.1 | 0 | 0 | 0.0 | 29.5 | 4 | 2 | 15.8 | 28.9 | 3 | 0 | 15.0 |
| Office Man | 31.1 | 28.8 | 30.8 | 0 | 0 | 0.0 | 30.4 | 11 | 4 | 39.5 | 32.5 | 3 | 1 | 20.0 |
| Purchasing Agent | 24.7 | 6.1 | 22.9 | 0 | 0 | 0.0 | 24.4 | 2 | 0 | 5.3 | 25.8 | 2 | 0 | 10.0 |
| Banker | 27.2 | 0.0 | 24.6 | 0 | 0 | 0.0 | 26.8 | 0 | 0 | 0.0 | 26.5 | 0 | 0 | 0.0 |
| Mortician | 28.7 | 7.6 | 25.9 | 0 | 0 | 0.0 | 29.6 | 3 | 1 | 10.5 | 28.1 | 0 | 1 | 5.0 |
| Salesmanager | 26.1 | 3.0 | 24.9 | 0 | 0 | 0.0 | 26.6 | 2 | 0 | 5.3 | 25.6 | 0 | 0 | 0.0 |
| Real Estate Salesman | 31.2 | 6.1 | 30.5 | 0 | 0 | 0.0 | 31.1 | 3 | 0 | 7.9 | 31.5 | 1 | 0 | 5.0 |
| Life Insurance Salesman | 28.3 | 10.6 | 26.5 | 1 | 0 | 12.5 | 29.2 | 5 | 0 | 13.2 | 27.5 | 1 | 0 | 5.0 |
| Advertising Man | 30.2 | 10.6 | 32.8 | 2 | 0 | 25.0 | 29.9 | 3 | 1 | 10.5 | 29.7 | 1 | 0 | 5.0 |
| Lawyer | 31.4 | 18.2 | 33.0 | 0 | 1 | 12.5 | 31.0 | 6 | 2 | 21.1 | 31.5 | 2 | 1 | 15.0 |
| Author-Journalist | 31.2 | 15.2 | 33.6 | 0 | 1 | 12.5 | 30.8 | 6 | 1 | 18.4 | 30.9 | 2 | 0 | 10.0 |
| President of Manufacturing Concern | 27.1 | 6.1 | 27.6 | 2 | 0 | 25.0 | 26.8 | 1 | 0 | 2.6 | 27.3 | 1 | 0 | 5.0 |

* Per cent of group making significant individual scores.

## TABLE 2

MEAN ALLPORT-VERNON SCORES AND THE SPREAD OF SIGNIFICANT INDIVIDUAL SCORES OF 66 MALE FIRST YEAR NEGRO MEDICAL STUDENTS FOR THE GROUP AS A WHOLE AND FOR EACH GRADE LEVEL

| | Entire group *N*=66 | | | *B* Grade students *N*=8 | | | *C* Grade students *N*=38 | | | *D* Grade students *N*=20 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean score | N SS* | % SS** | Mean score | N SS | % SS | Mean score | N SS | % SS | Mean score | N SS | % SS |
| Theoretical | 37.8 | 22 | 33.3% | 38.2 | 4 | 50% | 37.3 | 14 | 36.8% | 38.8 | 4 | 20% |
| Economic | 26.1 | 2 | 3.0 | 23.7 | 0 | 0 | 25.3 | 0 | 0 | 28.5 | 2 | 10 |
| Aesthetic | 23.5 | 1 | 1.5 | 23.6 | 0 | 0 | 23.3 | 0 | 0 | 23.9 | 1 | 5 |
| Social | 33.0 | 12 | 18.2 | 34.7 | 3 | 37.5 | 33.1 | 7 | 18.4 | 32.2 | 2 | 10 |
| Political | 28.2 | 5 | 7.6 | 26.8 | 0 | 0 | 27.7 | 1 | 2.6 | 29.6 | 4 | 20 |
| Religious | 33.0 | 12 | 18.2 | 33.0 | 2 | 25.0 | 33.3 | 8 | 21.1 | 32.3 | 2 | 10 |

* Number of significant individual scores.

** Per cent of group making significant individual scores.

Accountant, Office Man, Purchasing Agent, Banker, Mortician, Salesmanager, Real Estate Salesman; for *C* grade students from a high of 73.7 (Osteopath) to a low of 00 per cent (Banker); for *D* grade students from a high of 75 per cent (Osteopath) to a low of 00 per cent (Certified Public Accountant, Banker, Salesmanager).

7. Comparing *B* grade with *C* and *D* grade students in the matter of the number of significant individual vocational interest scores (*B*+ and *A*) for each of the four vocations in which *B* students as a group had a mean score of highly significant intensity (Physician, Osteopath, Mathematics-Physical Science Teacher, and Musician), one finds that the *B* grade group differs significantly from the *C* and *D* grade groups respectively in the number of highly significant individual interest scores only for Physician and Musician. For Physician the per cent of significant individual scores of the *B* grade group (87.5 per cent) is significantly higher ($t$=2.82) than that of the *C* grade students (47.4 per cent) and also significantly higher ($t$=2.31) than that of the *D* grade students (50 per cent). For Musician the per cent of significant individual interest scores of the *B* grade group (75 per cent) is significantly higher ($t$=2.15) than that of the *D* grade group (35 per cent) but not significantly different ($t$=1.29) from that of the *C* grade students (52.6 per cent).

In Table 2 are shown for the 66 first year medical students: *(a)* mean Allport-Vernon scores of the class as a whole in the six evaluative attitudes (Theoretical, Economic, Aesthetic, Social, Political, Religious); *(b)* the per cent of the class making on the Allport-Vernon significantly high individual scores (40 through 60) in each of the six evaluative attitudes; *(c)* the mean Allport-Vernon scores of the *B, C, D* grade groups respectively in each of the six evaluative attitudes; *(d)* the per cent of each grade group *(B, C, D* respectively) making significantly high individual Allport-Vernon scores (40 through 60) in each of the six evaluative attitudes.

In Table 2 the salient points appear to be the following:

1. In terms of mean Allport-Vernon scores of the class as a whole ($N$=66) for the six evaluative attitudes, there was no significant mean score (40 through 60) for any of the six evaluative attitudes. The per cent of the class as a whole ($N$=66) making significant individual scores (40 through 60) varied from a high of 33.3 per cent (Theoretical) to a low of 1.5 per cent (Aesthetic) with Social (18.2 per cent) and Religious (18.2 per cent) tied for second highest. The correlation between the mean scores and the per cent of significant individual scores for the six values is exactly +1.00.

2. Comparing the medical school grade levels *(B, C, D)* of these first year students as to scatter and intensity of evaluative attitudes, one notes *(a)*

that in terms of mean scores for each of the three grade levels there are at all grade levels no significant mean scores in any of the six evaluative attitudes; *(b)* that mean scores for the six values do not differ significantly from one grade level to another; *(c)* that the rank order of the mean scores of the *B* grade students namely, Theoretical, Social, Religious, Political, Economic, Aesthetic, does not differ from that of the *C* or *D* grade students save that the religious value ranks second place and the social value ranks third place in *C* and *D* students.

3. The *B* grade students do not differ significantly in the per cent of significant individual value scores from the *C* or *D* grade students in any value save Political where with $t=2.25$ the *D* grade group has a significantly higher per cent.

## C. Conclusions

The following conclusions are drawn from this study of 66 first year male Negro medical students:

1. The spread of vocational interests of significantly high intensity as seen in mean group scores extends to four vocations (Physician, Osteopath, Mathematics-Physical Science Teacher, Musician) for *B* grade students and only to two vocations (Osteopath and Mathematics-Physical Science Teacher) for *C* and *D* grade students.

2. In terms of significantly high individual scores the spread of vocational interests appears to be more extensive in the *C* and *D* grade students than in the *B* grade students.

3. Physician interest is found in significantly high intensity in 53 per cent of the 66 first year medical stutents and in as high as 87.5 per cent of the *B* grade students.

4. In this study all significant Strong scores, either individual or mean, fall mostly in *B*+ and less frequently just over the threshold of *A*, 60 being the highest individual score obtained by any one.

5. In terms of mean group scores no significantly intense evaluative attitudes were found at any of the three grade levels, the nearest approach to significant intensity being in the theoretical at all three grade levels (38.2, 37.3, and 38.8 respectively).

6. In terms of significantly high individual scores (Allport-Vernon) the spread of evaluative attitudes is more extensive in the lower grade students.

7. While no highly significant group scores were obtained, the few highly significant individual scores (Allport-Vernon) barely attain the threshold of significant intensity, 53 being the highest individual score obtained by any one.

## D. Discussion

That the vocational interest intensity of Negroes is significantly lower than that of whites has become a consistent finding with the Strong Vocational Interest Blank. The present study is no exception in its finding of low vocational interest intensity, particularly low physician interest, in male first year Negro medical students. Controversy turns not about whether the Negro's vocational interest intensity is lower or not, but about the reasons why it is lower, Hartshorn attributing it wholly to the effect of racial discrimination, Strong to racial segregation and other factors such as temperament, possibly lower intelligence, ambition (wishful thinking) without real interest. The present study would contribute to this controversy some evidence to the effect that in terms of significant individual scores the higher the scholastic ability of the Negro medical student the more intense is his physician interest and that the lower his scholastic ability the less intense is his physician interest and the more intense and/or diversified are his non-physician interests.

As we see it, it is in the dearth of high scholastic ability in this Negro group (out of 66 students no *A* and only eight *B* students) that the primary cause of the low intensity of physician interest of this group is lodged. On the other hand, racial segregation merely furnishes opportunity (segregated school, staff, hospital, etc.) for larger numbers of Negroes without significant physician interest or high scholastic ability to pursue medical studies.

## References

1. Hartshorn, H. H. Vocational interest patterns of Negro professional men. Unpublished Ph. D. thesis, August 1948, University of Minnesota Library.
2. Strong, E. K., Jr. Interests of Negroes and whites, *J. Soc. Psychol.,* 1952, **35**, 139-150.
3. Strong, E. K., Jr., & Tucker, A. C. The use of vocational interest scales in planning a medical career. Psychol. Monog., 1952, **66**, Serial No. 341.

*Department of Psychology*
*Howard University*
*Washington 1, D. C.*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 39-43.

# PERCEPTUAL ORGANIZATION IN A STUDY OF CREATIVITY*[1]

*Department of Psychology, University of Chicago*

MORRIS I. STEIN AND BERNARD MEER

The perception of good *Gestalten* is a function of the perceiver, his needs, defense mechanisms, and integration, and the characteristics of the stimulus-field. When the stimulus is ambiguous, the perceiver has to draw more heavily on his own resources than when the stimulus is structured. Consequently, it was hypothesized that those who have such resources available to them ("more creative" subjects) will develop more hypotheses and better *Gestalten* under varying conditions of ambiguity than those who may not have such resources available to them ("less creative" subjects).

Eighteen industrial research chemists who were expected to be creative and had the necessary facilities available to them for creative work served as subjects in this experiment.[2] Each of the subjects was rated by his superiors and colleagues on a global variable "creativity." The high group was rated significantly higher on this variable than the so-called lesser creative subjects.

The ambiguous stimuli in this experiment were the 10 standard Rorschach Cards. To vary the ambiguity of the stimuli, they were exposed by means of a Harvard tachistoscope at four different exposure levels—.01 of a second, .10 of a second, 3.0 seconds, and unlimited ("full") exposure. The subject was seated before the tachistoscope and instructed in the manner suggested by Klopfer and Kelley for traditional Rorschach administration, with one exception. Since each card was exposed at four different levels the subject was instructed not to make any conscious attempts to recall previous responses but if they did occur to him without intentional recall then he should verbalize his response.

Only subject's first responses to the cards were scored for the purpose of this experiment. These responses were first evaluated in terms of an 8-point rating scale including the following steps:

---

*Received in the Editorial Office on July 7, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Paper presented at the Mid-Western Psychological Association Meetings, Chicago, Illinois, May, 1953.

[2] This study was supported by a grant from the Research Division of Armour and Co., Chicago, Illinois.

1. *Autistic:* Those responses that fall into this category are of such a nature that there is little, if any, congruence, in the opinion of the raters, between the verbalized percept and the stimulus material.

2. *Rejection:* A Rorschach card was said to be rejected when the subject did not give a response to a card (e.g., "Nothing at all.") or if he suggested a response and then denied it (e.g., "It may be a brooch. No, it really doesn't remind me of anything.")

3. *Card Description:* Responses assigned to this category are those in which the subject described or named the physical aspects of the card but was unable to integrate them into a response that had definite shapes or form characteristics. Thus the subject might say, "I saw red, blue, and green," or, "It was all black." The response indicates that there is an "outthereness" quality about the card. There is, however, only a bare minimal interaction between the subject and the stimulus.

4. *Sensory:* This rating was reserved for those responses in which the subject utilized only the sensory aspects (or literal factors in Gibson's terms) of the cards. These factors were color, shading, and those responses that were determined solely by form where the form is of such a general nature that it could refer to most any type of stimulus (e.g., a "map" response that had no specificity).

5. *Sensory-form:* Responses that received this rating were those in which the sensory aspects of the card played a more important rôle than the form qualities of the card. Here is the first stage at which the controlling factors of the individual (if we may interpret the form qualities of a response as such) come into play. The sensory-form responses represent a more active attempt on the part of the perceiver to do something with his environment in a specific and socially communicable fashion than do the responses that were included in Categories 1-4.

6. *Fair-form:* In this category were included those responses in which there was a fair amount of agreement, in the opinion of the raters, between the responses and the stimuli. Form here is used in a general way to refer not only to "form" responses in the traditional Rorschach sense, but also to those responses that involve form plus the sensory aspects of the blots.

7. *Form dependent primarily on the demand characteristics of the stimuli:* This category was reserved for those responses that are scored as "populars" or are regarded as "near populars." This category reflects the fact that the perceiver's perceptions are, in general, congruent with those developed by the general community.

8. *Well-integrated form:* This rating was assigned to those responses in

which the perceiver provided sufficient evidence for his response and, in the opinion of the judges, the response was congruent with the stimulus. It usually involved the integration of achromatic or chromatic areas that were separated by white spaces or the integration of adjacent areas. A well-integrated response could also be assigned to a single area in which color or shading was utilized for a precise delineation of parts. Finally, a well-integrated response could be given to a popular area if a subject included something that was not so usual in his response and substantiated it with evidence. This rating suggests that a subject was capable of achieving good *Gestalten* and that, by inference, one may suggest that this reflected the integration of the individual personality. To be sure, even responses in this category as in the previous ones reflected some of the needs and interests of the subjects, but unlike some of the responses that were assigned to other categories, here the subject was able to present sufficient data to substantiate his percept.

The reliability of this scale had been previously established in another study.

Utilizing a previous population of 1,250 responses, a weighting system was then established which took into account the form-level rating of the response, and the frequency with which that specific rating appeared to a specific card at each of the exposure levels. Thus, for example, well-integrated responses to the most difficult cards (i.e., cards which the normative data indicated would not yield many well-integrated responses) at the shortest exposure level received the highest weights, while autistic responses and rejects at "full-exposure" received the greatest penalties. The range of weighted scores was between —3.5 and +4.0.

Having assigned a score to each response a total weighted score was established for each of the subjects. A bi-serial correlation between the weighted form-level score and the rating on the creativity variable was +.88, significant at the .01 level. Furthermore, analysis of the difference between mean weighted scores at each of the exposure levels indicated that high and low groups were significantly different from each other beyond the .01 level of confidence at all four exposure levels.

Further analysis of the data revealed that the sheer number of hypotheses suggested also differentiated the high from the low subjects. Looking at the data differently, we found that the high subjects rejected only 9 per cent of the cards while the low subjects rejected 31 per cent when exposure levels were collapsed. At the .01 exposure-level, the "highs" rejected 11 per cent and the "lows" 38 per cent, and at full-exposure, the respective figures were 1 per cent and 24 per cent.

The fact that the high subjects offered more hypotheses than the low subjects resulted in more autistic responses by the "highs" than the "lows" at the .01 exposure-level (31 per cent to 20 per cent). On the other hand, 51 per cent of the "highs'" hypotheses were rated fair or better at this exposure level while only 29 per cent of the "lows'" hypotheses achieved this rating.

As exposure-level became longer, the difference in autistic responses between the high and low subjects disappeared (8 per cent as against 4 per cent) while the difference in well-integrated responses between the two groups increased (41 per cent as against 16 per cent). The difference in rejects between the two groups remained relatively constant. At full-exposure, even when all information about the stimuli was available the low subjects still rejected 24 per cent of the stimuli. Of the remaining 76 per cent, almost one-half (32 per cent) of the hypotheses suggested were of the popular variety. Only 16 per cent of the responses at this exposure level achieved well-integrated forms. In contrast to this, the high subjects gave well-integrated responses to 41 per cent, populars to 29 per cent, and rejects to only one per cent of the stimuli.

The overall analysis showed the "highs" achieved significantly more well-integrated responses than the "lows" (.001 level of confidence). The question still remains whether these differences are due to a difference in intellectual ability and/or differences in personality factors, such as anxiety, defensiveness, etc., which result in over-cautiousness and over-criticalness which would tend to inhibit the development of hypotheses. We found significant correlations between the Wechsler-Bellevue Full Scale Scores and creativity and between the Rorschach scores and the Wechsler-Bellevue Full Scale Scores. Partialling out the effects of intelligence it was found that most of the variance contributing to the relationship between the Rorschach and creativity is due to factors other than intelligence, although the latter still is of some importance here.

Our data provide some additional evidence that personality factors are probably most crucial. This stems from the fact that if we hold the number of rejects constant we find that the two groups of subjects do not differ significantly from each other in terms of the number of well-integrated responses. This strongly suggests that the real difference between our high and low creative individuals may turn out to be a function of defensiveness or over-criticalness which inhibits the generation and communication of hypotheses.

In conclusion it may be said that "more creative" research chemists show a greater freedom in the offering of perceptual hypotheses and in the level of organization of their responses, than do their colleagues who are rated as "less creative." This appears to be a function of personality factors which we hope to delineate more specifically in the future.

*Department of Psychology*
*University of Chicago*
*Chicago 37, Illinois*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 45-50.

# SIMILARITY IN TEACHER AND PUPIL PERSONALITY*

*Department of Psychology, St. Francis College*

SISTER MARY AMATORA, O.S.F.

## A. THE PROBLEM

Prominence of personality as the number one quality of the teacher is no longer a disputed topic. That the personality of the pupil is influenced by the personality of the teacher is likewise maintained by many educators today. A limited amount of vital research in this area has appeared in recent years, yet much remains to be studied before any final answers will be forthcoming.

Symonds (16), who has done considerable work in this area, states emphatically, "If better teachers are to be employed in our schools, more attention must be paid to personality factors in their selection."

In Robbins' (14) study, "ability to make the course interesting" received first place, above "teacher's knowledge of subject matter." The same study showed the most undesirable trait to be the teacher's failure to treat the student as an adult. These again show the high position accorded personality elements in the estimation of pupils.

Anderson's (9) classic studies on the effects of dominative and integrative behavior of the teacher upon pupil behavior have shown quite conclusively that the direction is from teacher to pupil, but never vice versa. Herrick's (12) observations are in the same direction.

Cook and Leeds (11) find that "the attitude of individual teachers toward pupils is significantly related to the pupils' attitudes toward the teachers."

Tiedeman's (17) investigation of junior high school pupils likewise brought to the fore the personality qualities of the teachers as uppermost in pupil estimation. Reference to other studies (4, 13, 3, 6, 15) shows similar trends.

## B. EXPERIMENTAL DESIGN

The present study seeks to investigate the relationships, if any, between the personalities of pupils and the personalities of their respective teachers on certain specified components of personality.

### 1. *Method*

Of the various types of measuring instruments for studying personality, such as the objective test, the life record data, and the questionnaire and

---

*Received in the Editorial Office on July 11, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

rating scale technique, the latter method was chosen for the present experiment. As the inventory type of questionnaire is usually limited to self-ratings, a further element of objectivity was achieved by using the rating scale, on which a number of independent judgments could be secured for each individual pupil and teacher in the study.

*a. The scale.* The Child Personality Scale (1, 4, 7) and its comparable adult form (2, 5, 8) prepared by the writer for earlier investigations were used. These instruments provide for the rating of both children and adults on 22 elements of personality. Each component of the personality is treated as a separate scale; each individual is rated separately on each scale on a 10-point continuum. The ends of the scales are staggered so as to avoid any possible halo effect.

Interpretation of each end of the scale is given by descriptive words and/or phrases, as is the mid-point or point of indifference signifying the average or "most people."

The 22 elements of personality included in the study are: Kindness, sympathy and thoughtfulness of others; dependability; politeness and courtesy; neatness and cleanliness; sociability; nervousness or calmness; punctuality in meeting appointments; honesty and fairness; patient and good-natured or easily angered; interested in few or in many things; quiet or boisterous; sense of humor; good-sportsmanship; coöperative in working with others; intelligent; energetic and peppy; popularity; religiousness; generosity; persistence in sticking to a task; optimistic or pessimistic; and, boring or entertaining.

*b. The subjects.* Subjects used in the experiment were 100 teachers and their pupils of grades four through eight in both public and private schools, in both city and rural schools.

*c. Procedure.* Each teacher was rated on each of the 22 scales by three or four fellow teachers in the same school. The children were rated by eight classmates on each of the 22 scales. For the second part of the experiment ratings were secured for each child on each scale by three or four teachers.

## C. Analyses of Data

When all papers were completed, data were analyzed separately for teachers and for pupils, in each of the two parts of the study.

### 1. *The Teachers' Scores*

The three or four ratings each teacher received on each scale were averaged separately for each teacher on each scale. These became the individual teacher's raw score ratings on the specified scales.

### 2. *The Pupils' Scores*

For the pupils' scores there were two separate sets of data. Each was treated separately throughout the experiment: (*a*) The ratings each child received by the eight classmates were averaged on each of the twenty-two scales. This provided the first set of raw scores for each pupil on each scale. From these, means were computed for each class on each scale. (*b*) Each child's ratings by three or four teachers on each scale were averaged for the second set of data on the pupils' personalities. From these, the second set of class means was computed for each of the 22 scales.

### 3. *Correlations*

Pearsonian *r*'s were computed separately for the two sets of data. For each of these, the mean pupil personality score on each scale was plotted against the average score for the teacher of that respective class. Thus were separate measures of correlation established between teacher and pupil personality on each of the 22 variables measured in the present study. The *N* for each *r* is 100 in each of the two sets of data.

## D. RESULTS AND INTERPRETATIONS

As the object of this study is to determine the relationships, if any, between 22 specified components of personality in teacher and in pupil, two separate analyses are made. The first of these is based on the personalities of the teachers and their pupils, both as judged by teachers; and, the second compares the personalities of teachers as judged by teachers with the personalities of pupils as judged by pupils.

### 1. *When Teachers do the Rating*

When the children's personalities in terms of teacher judgments are correlated with the personalities of their respective teachers in terms of fellow-teachers' ratings, the *r*'s given in Table 1 are achieved. These, arranged in rank order from highest to lowest, reveal a considerable degree of relationship between teacher and pupil personality on most of the traits measured.

The *r*'s corrected for attenuation range from .46 to .11 with a median *r* of .275. According to Fisher's test of significance[1] all *r*'s above .197 are significant at the 5 per cent level of confidence, and all *r*'s above .256 are significant at the one per cent level. Hence, at the one per cent level of confidence there is a statistical significance in the similarity between teacher and pupil personality on the first 15 scales listed in this table. On four more scales there is statistical significance at the 5 per cent level. The only three scales on which the relationship between teacher personality and pupil personality is below the 5 per cent level of confidence are those for sense of humor, good-sportsmanship, and religiousness. However, all correlations are positive, even on these low *r*'s.

---

[1] *Psychometric Methods* by J. P. Guilford (p. 335 ff.).

## 2. *When Children do the Rating*

To corroborate the above findings, a second set of data was analyzed in the same manner as was the first set. Pearsonian *r*'s were again computed for the 22 separate scales. This time the scores for the pupils' personalities as judged by their classmates were plotted against the scores of the teacher personality as rated by fellow-teachers.

The results are given in Table 2. It will be noted that, though slightly lower on some of the scales, the *r*'s of this table are not greatly different from those in Table 1. The present *r*'s range from .41 to .11 with a median *r* of .22. Here again, for the two variables and 98 degrees of freedom, all *r*'s at and above .197 are significant at the 5 per cent level, while all *r*'s at .256 or above, are statistically significant at the one per cent level of confidence.

In this table the relationships between teacher and pupil personality are statistically significant at the one per cent level on 10 of the 22 scales, and at the 5 per cent level on five more of the scales. Again, all correlations are positive. The findings of this second set of data corroborate the results of the first set of data in the present experiment.

TABLE 1

CORRELATION BETWEEN PUPILS' AND TEACHERS' PERSONALITIES I

| Rank | Personality variables | $r_{\infty w}$ | $r \pm SE$ | | Rank | Personality variables | $r_{\infty w}$ | $r \pm SE$ | |
|---|---|---|---|---|---|---|---|---|---|
| 1. | Energy and pep | .46 | .38 | .08 | 12. | Boisterous-quietness | .27 | .24 | .09 |
| 2. | Sociability | .43 | .34 | .09 | 13. | Thoughtfulness | .27 | .24 | .09 |
| 3. | Nervous-calmness | .41 | .33 | .09 | 14. | Boring-entertaining | .26 | .23 | .09 |
| 4. | Generosity | .40 | .31 | .09 | 15. | Neatness | .25 | .22 | .09 |
| 5. | Cooperation | .34 | .28 | .09 | 16. | Interests | .23 | .20 | .09 |
| 6. | Persistence | .33 | .28 | .09 | 17. | Courtesy | .23 | .19 | .10 |
| 7. | Disposition | .32 | .26 | .09 | 18. | Honesty | .20 | .17 | .10 |
| 8. | Punctuality | .31 | .26 | .09 | 19. | Dependability | .20 | .16 | .10 |
| 9. | Intelligence | .29 | .25 | .09 | 20. | Sense of humor | .16 | .13 | .10 |
| 10. | Popularity | .29 | .25 | .09 | 21. | Good-sportsmanship | .16 | .12 | .10 |
| 11. | Patience | .28 | .24 | .09 | 22. | Religiousness | .11 | .06 | .11 |

TABLE 2

CORRELATION BETWEEN PUPILS' AND TEACHERS' PERSONALITIES II

| Rank | Personality variables | $r_{\infty w}$ | $r \pm SE$ | | Rank | Personality variables | $r_{\infty w}$ | $r \pm SE$ | |
|---|---|---|---|---|---|---|---|---|---|
| 1. | Energy and pep | .41 | .34 | .09 | 12. | Neatness | .21 | .18 | .10 |
| 2. | Sociability | .38 | .31 | .09 | 13. | Intelligence | .20 | .18 | .10 |
| 3. | Nervous-calmness | .37 | .30 | .09 | 14. | Persistence | .20 | .17 | .10 |
| 4. | Popularity | .34 | .25 | .09 | 15. | Boisterous-quietness | .20 | .17 | .10 |
| 5. | Generosity | .32 | .25 | .09 | 16. | Interests | .17 | .15 | .10 |
| 6. | Thoughtfulness | .31 | .24 | .09 | 17. | Punctuality | .17 | .14 | .10 |
| 7. | Patience | .30 | .24 | .09 | 18. | Honesty | .16 | .13 | .10 |
| 8. | Courtesy | .28 | .23 | .09 | 19. | Dependability | .15 | .12 | .10 |
| 9. | Cooperation | .27 | .22 | .09 | 20. | Good-sportsmanship | .14 | .11 | .10 |
| 10. | Disposition | .27 | .22 | .09 | 21. | Religiousness | .13 | .09 | .10 |
| 11. | Boring-entertaining | .23 | .19 | .10 | 22. | Sense of humor | .11 | .09 | .10 |

## E. SUMMARY AND CONCLUSIONS

A study of the personalities of 100 classroom teachers and their pupils revealed positive relationships on all the 22 elements of personality measured.

Results were analyzed from two separate sets of data, which corroborated each other. On more than half the scales the similarity between teacher and pupil personality showed statistical significance at the one per cent level of confidence. For this sample of teachers and pupils there is less than one chance in a hundred that these *r*'s could have occurred if there were no true correlation. On about a fourth or more of the scales the similarity was verified within the limits of the 5 per cent level; that is, there is less than five chances in a hundred that these *r*'s could have arisen if there were no true correlation.

The most important finding of this experiment is the complete absence of all negative correlations. On every element of personality measured in this study there is a positive relationship between teacher and pupil personality.

Statistical data do not yield answers relative to *causal* relationships. Facts only are reported. Yet, other studies mentioned earlier in this article suggest such between pupil and teacher. However, if the findings of this study be true generally, then it is of vital importance to the development of wholesome personality in the children, that they have teachers who possess well-adjusted personalities.

## REFERENCES

1. AMATORA, S. M. A twenty-two trait personality scale. *J. of Psychol.*, 1944, **18**, 3-8.
2. ————. Boys' personality appraisals differentiate teacher groups. *Sch. & Soc.*, 1952, **76**, 184-187.
3. ————. Can elementary school children discriminate certain traits in their teachers? *Child Devel.*, 1952, **23**, 75-80.
4. ————. Child Personality Scale. Cincinnati: Gregory, 1951.
5. ————. Comparability of child and adult personality scales. *J. Educ. Psychol.*, 1944, **35**, 309-313.
6. ————. Diagnostic Teacher-Rating Scale. Cincinnati: Gregory, 1952.
7. ————. Factor analysis of children's personality scale. *J. of Psychol.*, 1944, **18**, 197-201.
8. ————. The education factor in personality appraisal. *J. Exp. Psychol.*, 1953, **22**, 271-275.
9. ANDERSON, H. H., BREWER, J. E., & REED, M. F. Studies of teachers' classroom personalities: III. Appl. Psychol. Monog., 1946, No. 11.
10. BAKER, H. J. Integration of character and personality. *Nat. Educ. Ass. Proc.*, 1940, 631-638.
11. COOK, W. W., & LEEDS, C. H. Measuring the teaching personality. *Educ. Psychol. Measmt.*, 1947, **7**, 399-410.
12. HERRICK, V. E. Teachers' classroom personalities. *Elem. Sch. J.*, 1945, **46**, 126-129.
13. MATTHEWS, R. A. Effects of teacher personality on the emotional life of the child. (*In Roads to the Future.*) *Sch. Educ. Bull.*, 443-453.
14. ROBBINS, F. G. Student reactions to teacher personality traits. *Educ. Admin. & Suprv.*, 1944, **30**, 241-246.
15. SNYDER, W. U. Do teachers cause maladjustment? *J. Except. Child.*, 1947, **14**, 40-46, 73-78.
16. SYMONDS, P. M. Evaluation of teacher personality. *Teach. Coll. Rec.*, 1946, **48**, 21-34.
17. TIEDEMAN, S. C. A study of pupil-teacher relationship. *J. Educ. Res.*, 1942, **35** 657-664.

*St. Elizabeth Hospital*
*Lafayette, Indiana*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 51-58.

# RORSCHACH STATISTICS ON A GROUP OF 136 NORMAL MEN*

*Urban Life Research Institute, Tulane University*

BARBARA WEDEMEYER[1]

## A. INTRODUCTION

This paper presents Rorschach statistics on a group of 136 men who were serving in the U. S. Navy at the time of the investigation. The subjects represent a group not usually tested in any great number: normal men of average intelligence. The majority of normative reports in the literature have been based on data from psychiatric patients or from normal adults in the high average to superior intelligence range.

The men used in this study were assigned for testing without regard for personal preference. This method of selection has some advantages, since it is conceivable that there is a difference in personality (and therefore in Rorschach pattern) between persons who volunteer for psychological testing and those who do not. Nonvoluntary participation creates new problems, however. The subjects' lack of motivation on the Rorschach, and on the other tests given, was very apparent. They were not interested in the tests, and in most cases, did the minimum necessary to comply with instructions.

The "normality" of the group may be questioned. All the men seen were normal in the sense that they functioned in society and had no mental illness requiring hospitalization. However, while a few men may have joined the Navy only to avoid the draft, it is important to note that most of them had voluntarily chosen the secure, sheltered life of an enlisted man in the regular Navy; a life, on the whole, relatively free of personal responsibility. The men as a group, with certain individual exceptions, impressed the psychologist as being somewhat passive, apathetic, and lacking in initiative and drive. This impression was corroborated by psychiatrists who interviewed the men.

---

*Received in the Editorial Office on July 13, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is based upon data gathered as a part of a research project under the direction of John H. Rohrer and supported by the Neuropsychiatry Branch, Professional Division, Bureau of Medicine and Surgery, Department of Navy through contract Nonr 475(01) between the Office of Naval Research and Tulane University. This paper is a technical report made under that contract. The opinions expressed herein are those of the writer and are not necessarily those of the sponsoring Agency, the Department of Navy.

While such speculation is interesting and possibly fruitful, it is not the concern of the present report, which consists of a non-interpretive presentation of Rorschach statistics.

## B. Description of Subjects

The *S*s in this investigation comprised a very homogeneous group. They were all Navy personnel below the Chief Petty Officer rating, and the great majority were Regular Navy. A breakdown of the ratings of the 122 men for whom this data was available is shown in Table 1. Virtually all of the men holding those ratings and on duty at the base at the time of the experiment were tested.

TABLE 1
Naval Rating of Subjects

| Rating | N |
|---|---|
| Seaman Apprentice | 21 |
| Seaman | 40 |
| Petty Officer Third Class | 23 |
| Petty Officer Second Class | 21 |
| Petty Officer First Class | 17 |
| No data available | 14 |
| Total N | 136 |

The mean age of the *S*s at last birthday was 23.7 years, with a *SD* of 3.92 and a range from 17 to 39 years. Eighty-seven per cent of the subjects were in the range from 19 to 28 years. Navy General Classification Test scores were available for 97 men. This test has a mean of 50 and a *SD* of 10 for the normal population. The mean for the present group was 51.79, *SD* 9.56, and the range of scores 28 to 82. All but six of the scores (94 per cent) fell in the range from 36 to 67. The group as a whole can be taken to represent the middle portion of the normal intelligence range, if one assumes the *GCT* to be a valid measure of intellectual capacity.

This finding is of particular interest in view of the paucity of normative reports on subjects of average intelligence. Gardner (5) reports *IQ*'s in the average range for 100 nurses. Cass and McReynolds (3) present Rorschach data on a group of 104 adults whose occupations are representative of the general adult population. Beck's report (2) on the normal personality contains no intelligence test data. However, he does report the highest school grade reached by each of his subjects. A comparison of these figures with those derived from US census data (7) is made in Table 2. The percentage of Beck's subjects who have had some high school education is approximately two and one-half times greater than the percentage for the total population, while the percentage who have had no more than grammar school education

TABLE 2
EDUCATION BY HIGHEST GRADE REACHED

| | Beck's Group (N=154) | US Census |
|---|---|---|
| College | 13.0% | 10.0% |
| High School | 74.4 | 29.1 |
| Grammar School | 12.3 | 55.8 |
| No schooling | —— | 5.1 |

is correspondingly less. The percentage of college trained people is approximately the same. Since there is a correlation between educational level attained and measured intelligence we can infer that Beck's group is above average in intelligence. While undoubtedly most of his subjects do fall within "the middle two-thirds of the population in intelligence" (2, p. 244), they probably cluster at the upper end of this range.

### C. TESTING PROCEDURE

The subjects were assigned to report for testing as a part of a project concerned with the relationship of perceptual learning, and personality variables to leadership. The Rorschach was introduced as one of several tests to be given. It was not mentioned by name and only four subjects reported ever having seen it before.

The Rorschach was administered following standard procedure. Each man was seen alone in a private office. A brief explanation of the procedure was given, every effort being made to relax and interest the subject. His chair was placed so that he would sit with his back to the psychologist. However, in an effort to secure maximum spontaneity, no comment was made if he turned in his chair during the administration of the test. The following instructions were given:

> I am going to show you a series of pictures which were made with splotches of ink, so they don't look like anything in particular. Different people see all sorts of different things in them; and we're interested in the kinds of things that you can see. Now, there are no right or wrong answers, and no one particular thing to look for. I want you to look at the cards and tell me what you can see in them, what they look like to you.

On the first three cards the subject was encouraged to be more productive, usually with the comment, "can you see anything else?" The usual inquiry was made, followed by a "testing of the limits" (6). A stop watch was used for timing.

Three psychologists administered the Rorschach. Ninety-two of the records were secured by one person, who also scored all the protocols. The psychologists were women; all had an MS degree and at least a year's clinical experience which included Rorschach testing. There was no difference in the kind of record obtained by the various examiners. Emphasis is placed on procedure and on the qualifications of the psychologists because of the very atypical, meager records obtained from a presumably normal group.

## D. Test Findings

Results will be presented in every case for 136 subjects. One hundred and thirty-eight men were seen originally. One was eliminated because he was a Filipino with a poor command of English, and the other because his total of 61 responses was almost twice as great as that of the man having the next highest total (34 responses). It was felt that his inclusion would markedly distort the data.

Some explanation of the scoring system is necessary. The psychologist who scored all the tests was trained in the Klopfer technique. However, in those instances where a sufficient body of statistical data existed, these data were used as the basis for scoring. The statistics used were those presented by Beck in his *Rorschach's Test,* Vol. I (1), as revised by his article on the normal personality (2). The scoring of D and Dd, F+ and P followed his tables strictly. Scoring of W, M, and color followed the criteria stated in Beck's book (1). It was not felt that this scoring of the latter categories would differ markedly from Klopfer's, except possibly for W. A few responses scored D, after Beck, might have been scored W by Klopfer and included by him in the W classification. F was scored by both Klopfer's and Beck's criteria and is discussed below. The remaining determinants, FM, m, k, K, FK, c, C′, and the content categories, were scored following Klopfer (6). Unless otherwise stated, the statistics treat only main responses.

Table 3 presents the M, Mdn, *SD,* and range of scores for each Rorschach variable. The distribution of scores was plotted for each category and in almost every case, the curve of the distribution was quite smooth. The type of curve, direction of skewness, and mode or modes are described in Table 3. In general, Table 3 is self-explanatory. Ambiguous points, marked with an asterisk, have been clarified below.

### 1. *Location Categories*

Responses scored *S main* were made primarily to the white space and represented a true figure-ground reversal. Responses scored *S add* were those used as elaboration (eyes, mouth, etc.) of a main response to the blot figure. Because of the small number of responses to each card, no measure of sequence could be made.

## TABLE 3
### Statistical Breakdown of Rorschach Data

| Scoring Category | Mean | Meanσ | Median | Mdn.σ | Range | Mode | Type of Curve |
|---|---|---|---|---|---|---|---|
| 1. Location Categories | | | | | | | |
| R Response total | 13.654 | 6.709 | 12 | 8.406 | 3-34 | 8,11 | Positively skewed |
| W Whole responses | 4.206 | 2.609 | 4 | 3.269 | 0-11 | 2 | Positively skewed |
| D Detail responses | 7.919 | 5.271 | 7 | 6.605 | 0-25 | 6 | Positively skewed |
| Dd Rare detail responses | 1.368 | 1.620 | 1 | 2.030 | 0-7 | 0 | J shaped |
| * S main | .162 | .545 | 0 | .683 | 0-3 | 0 | J shaped |
| * S additional | .794 | .924 | .5 | 1.158 | 0-3 | 0 | J shaped |
| 2. Determinant Categories | | | | | | | |
| M Human movement | .971 | 1.084 | 1 | 1.358 | 0-6 | 0 | J shaped |
| FM Animal movement | 2.324 | 1.785 | 2 | 2.237 | 0-8 | 2 | Positively skewed |
| * m Minor movement | .147 | .536 | 0 | .672 | 0-4 | 0 | J shaped |
| K +k Shading | .309 | .562 | 0 | .704 | 0-2 | 0 | J shaped |
| FK Vista | .287 | .640 | 0 | .802 | 0-3 | 0 | J shaped |
| * F Klopfer | 6.882 | 4.352 | 6 | 5.453 | 1-25 | 5 | Positively skewed |
| * F Beck | 9.331 | 5.263 | 8 | 6.595 | 2-26 | 5 | Positively skewed |
| c Texture | .846 | .938 | 1 | 1.175 | 0-4 | 0 | J shaped |
| C' Achromatic color | .419 | .733 | 0 | .918 | 0-3 | 0 | J shaped |
| FC Form-color | .691 | 1.003 | 0 | 1.257 | 0-4 | 0 | J shaped |
| CF Color-form | .537 | .830 | 0 | 1.040 | 0-3 | 0 | J shaped |
| C Color | .235 | .730 | 0 | .915 | 0-5 | 0 | J shaped |
| 3. Content categories | | | | | | | |
| H Humans | 1.147 | 1.147 | 1 | 1.437 | 0-6 | 1 | Positively skewed |
| Hd Human detail | .846 | 1.282 | 0 | 1.606 | 0-6 | 0 | J shaped |
| A Animals | 5.603 | 2.722 | 5 | 3.411 | 1-15 | 6 | Positively skewed |
| Ad Animal detail | 1.397 | 1.879 | 1 | 2.354 | 0-14 | 0 | J shaped |
| At Anatomy | .676 | 1.083 | 0 | 1.357 | 0-8 | 0 | J shaped |
| Obj Objects | .346 | .679 | 0 | .851 | 0-4 | 0 | J shaped |

| Scoring Category | Mean | Meanσ | Median | Mdn.σ | Range | Mode | Type of Curve |
|---|---|---|---|---|---|---|---|
| 3. Content categories con't. | | | | | | | |
| Pl Plants | .941 | 1.264 | 0 | 1.584 | 0-6 | 0 | J shaped |
| Nat Nature | .368 | .694 | 0 | .870 | 0-5 | 0 | J shaped |
| Geo Geography | .390 | .858 | 0 | 1.075 | 0-4 | 0 | J shaped |
| * No. of content categories | 6.897 | 2.259 | 7 | 2.830 | 2-14 | 7 | Normal |
| 4. Ratios, Percentages, etc. | | | | | | | |
| Z Organization (Beck) | 20.243 | 11.290 | 18 | 14.146 | 0-47.5 | 18 | Positively skewed |
| ΣC $\frac{FC + 2\ CF + 3\ C}{2}$ | 1.224 | 1.591 | .5 | 1.994 | 0-7.0 | 0 | J shaped |
| A o/o A + Ad/R | 53.346 | 16.019 | 53 | 20.072 | 19-100 | 43 | Normal |
| * F o/o Klopfer | 51.000 | 18.616 | 50 | 23.326 | 5-100 | 48 | Normal |
| * F o/o Beck (Lambda) | 69.132 | 17.615 | 71.5 | 22.072 | 20-100 | 73 | Negatively skewed |
| * F+ o/o Beck | 85.559 | 13.238 | 86 | 16.587 | 33-100 | 100 | J shaped |
| P popular responses | 4.757 | 2.066 | 5 | 2.589 | 0-11 | 4 | Normal |
| O Original responses | 2.713 | 2.525 | 2 | 3.164 | 0-11 | 1 | Positively skewed |
| * Time per response (sec.) | 29.824 | 13.575 | 27.5 | 17.009 | 11-96 | 24 | Positively skewed |
| * Reaction time, color | 18.596 | 13.511 | 17 | 16.929 | 4-97 | 11 | Positively skewed |
| * Reaction time, shading | 17.213 | 10.554 | 15 | 13.224 | 4-62 | 11 | Positively skewed |
| R VIII-X/Total R x 100 | 34.235 | 11.963 | 33 | 14.990 | 0-67 | 37 | Normal |
| R VIII-X/R I-VII | .575 | .323 | .500 | .405 | 0-2.5 | .5 | Normal |
| * Elizur A | 5.051 | 3.966 | 4 | 4.969 | 0-22 | 4 | Positively skewed |
| * Elizur H | 1.493 | 1.814 | 1 | 2.273 | 0-9 | 0 | J shaped |
| * Number of rejections | 1.640 | 1.888 | 1 | 2.366 | 0-7 | 0 | J shaped |

* Refer to text

### 2. *Determinant Categories*

F was scored two ways. *F Klopfer* was scored following Klopfer's criteria (6, pp. 108-157). The records were then rescored according to Beck's criteria for F (1, pp. 155-190). This scoring was labeled *F Beck*. Essentially, this response includes the categories scored by Klopfer as F, FM, and Fm.

### 3. *Content Categories*

The content categories other than those shown contained so few entries that it is meaningless to include them. *Number of content categories* refers to the number of different categories used. In an effort to increase N, both main and additional content responses were counted. The number of possible categories includes the 19 categories listed on the Klopfer record blank, plus one additional category for miscellaneous responses, such as clothing, map, explosion, etc. Thus, a maximum score of 20 is possible.

### 4. *Ratios, Percentages, etc.*

*F % Klopfer* was computed by dividing the total number of F Klopfer responses by R, and multiplying by 100.

*F % Beck* was computed by summing the total number of responses scored *F Beck* (whether F+ or F—), dividing by R, and multiplying by 100. The formula is that used by Beck to compute his lambda index (2, pp. 277-278).

All F responses were scored either + or —. Where the response did not appear in Beck's tables for F+ and F—, decision was made by the examiner on the basis of similarity to other responses scored F+ or F—, or on the basis of subjective impression of the degree of correspondence between the blot area and content of the response.

*Time per response* (T/R) was obtained by dividing the total time by the number of responses. If a card was rejected, the time for this card was not included in the total time.

*RT Col* refers to the mean reaction time to the five colored cards. Reaction time was defined as the time between presentation of the card and initial scorable response. This held even if the response was later rejected in the inquiry. *RT Sh* was similarly computed on the five black cards. All time scores are given in seconds.

The content was scored for anxiety and hostility according to Elizur's criteria (4). His method essentially involves weighting responses one or two points when they symbolically or openly express anxiety or hostility. The sum of the weightings for anxiety is referred to as *Elizur A*, and the sum of the weightings for hostility as *Elizur H*.

Finally, number of rejections was tabulated. This category does not include cases where the subject responded to a card during the initial performance but rejected this response during the inquiry.

The number of times each card was rejected is tabulated in Table 4.

Table 4
NUMBER OF REJECTIONS PER CARD

| Card Number | Times Rejected |
|---|---|
| IX | 42 |
| VII | 34 |
| VI | 33 |
| IV | 32 |
| X | 27 |
| II | 23 |
| III | 11 |
| I | 10 |
| VIII | 7 |
| V | 6 |

As would be expected, Card IX was the most frequently rejected. Following were the three heavily shaded cards, Cards VII, VI, and IV. Card X, with its confusion of colors and shapes, and Card II, the first of the color cards, were next. Cards III, I, VIII, and V apparently presented less of a problem to the subjects. Card III was rejected only one-fourth as often as Card IX, and half as often as Card II.

## E. SUMMARY

Rorschach data have been reported for 136 normal men of average intelligence, who were serving in the U.S. Navy at the time of investigation. The subjects, whose mean age was 23.7 years, were all enlisted men below the rank of Chief Petty Officer

The body of the paper comprised a statistical breakdown of the Rorschach data. The M, Mdn, *SD*, range, mode, and type of frequency curve were given for each of the usual Rorschach categories and for Elizur's (4) anxiety and hostility scores. A tabulation was made of the number of times each card was rejected.

The test records obtained can be characterized as meager and atypical of published norms on "normal" groups.

## REFERENCES

1. BECK, S. J. Rorschach's test. Vol. 1. *Basic processes.* New York: Grune & Stratton, 1944.
2. BECK, S. J., *et al.* The normal personality as projected in Rorschach test. *J. of Psychol.,* 1950, **30**, 241-298.
3. CASS, W. A., & MCREYNOLDS, P. A contribution to Rorschach norms. *J. Consult. Psychol.,* 1951, **15**, 178-184.
4. ELIZUR, A. Content analysis of the Rorschach with regard to anxiety and hostility. *Rorschach Res. Exch.,* 1949, **13**, 247-284.
5. GARDNER, G. E. Rorschach test replies and results in 100 normal adults of average *I.Q. Am. J. Orthopsychiat.,* 1936, **6**, 32-60.
6. KLOPFER, B., & KELLEY, D. M. The Rorschach Technique. New York: World Book, 1946.
7. Statistical Abstract of the U. S. 1946. U. S. Dept. Commerce, Bureau of the Census. Washington, D. C.: U. S. Government Printing Office.

*Urban Life Research Institute*
*Tulane University*
*New Orleans 18, Louisiana*

Published as a separate and in *The Journal of Psychology,* 1954, **37,** 59-64.

# FURTHER REMARKS ON THE HYPOTHETICAL CONSTRUCT*

*Department of Psychology, Hofstra College*

---

ROBERT PLUTCHIK

---

Since the MacCorquodale and Meehl paper appeared (6), which distinguished intervening variables from hypothetical constructs, there has been time for a reconsideration of the problem. Tolman (12), for example, has decided to give up intervening variables for hypothetical constructs; Skinner (10), by largely relinquishing the notion of the "reflex reserve," has retreated even further from such hypothetical terms. Krech (5) has taken up the cudgels on one side and Kessen and Kimble (4) on the other.

It seems to the present writer, however, that all the fire and enthusiasm on either side is not only one-sided, but misdirected, if we consider the general aim of science, which is to coördinate as large a number of facts in the simplest framework as is possible. This paper will attempt to show that the proposed distinction between hypothetical constructs and intervening variables, with the exception of the particular terms used, is not new to science but is a relatively late acquisition for psychology; and also, that the distinction, even if valid, does not enable the scientific worker to decide in advance which kind of theoretical term will be most fruitful. The justification for these criticisms will become clear as the historical context and present function of such constructs are traced.

The basic distinction which Meehl and MacCorquodale made is between terms which are simply summary words for a group of observations (intervening variables) and terms which refer to events, processes, or entities "not among the observed" (hypothetical constructs).

This is related to, but not identical with, distinctions which have been suggested in the past. For example, a physicist named Rankine (8) published a book in 1881 in which he made a distinction between an "abstractive" theory (such as Newton's stellar theory of mechanics) and an "hypothetical" theory (such as the kinetic theory of gases). The abstractive theory deals with certain selected features of macroscopic objects, such as velocity, position, mass, temperature, etc., and establishes relations between them. In a hypothetical theory a correspondence is established between an experimentally defined concept like temperature and a complex of ideas relating to assumed but unobserved events, like the mean kinetic energy of molecules.

---

*Received in the Editorial Office on July 16, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

Again a related but not identical type of distinction was suggested by the early German philosopher Wilhelm Dilthey (3). Dilthey distinguished between a "descriptive" and an "explanatory" science. The former is one whose terms and laws are found by empirical analysis, or close examination of what is actually given in experience. Sciences like astronomy and geology would tend to be descriptive. An "explanatory" science is one which takes its units and laws from a methodological assumption which determines their general nature beforehand. Atomic physics, which is such a theory, is based upon the notion of a world of imperceptible atoms possessing only primary qualities. This presumably explains the world of perception but does not actually describe it.

A similar but again somewhat different analysis was presented by the chemist and philosopher of science, E. Meyerson (7). He distinguished between a principle of causality and a principle of legality. When someone invokes a causal explanation he generally refers to some process involving unobserved entities which have certain invariant aspects. For example, as the pressure of a gas is made to increase, the total energy of the system remains unchanged or invariant, and the assumed atoms of the gas retain their identity, even though they may move more rapidly. The causal explanation is designed to relate the later and earlier stages of any changing process or system by means of something which doesn't change (atoms, light corpuscles, total energy or momentum, etc.).

A legal explanation would simply require a statement of the empirical rule which describes the change of an observed event with time, or in other words, a statement of the law of which the observed phenomenon is a special case. To explain the fall of a particular apple by saying that "all objects fall if left unsuspended" would invoke the principle of legality; to explain the fall of an object by assuming certain unobserved bodies acting by contact, or by attributing some particular property to the ether, as many physicists attempted to do in the nineteenth century, would provide a "causal" explanation.

These examples, which are by no means exhaustive, suggest that this kind of a distinction does in fact correspond to a real aspect of the process of science. One major philosophical development, at least, has attempted to indicate the path that science should follow. The positivists have always argued in favor of an "abstractive" or "descriptive" science governed by "legal" explanations and "intervening variables."

One might argue the legitimacy of the distinctions drawn above between observed and unobserved entities. Many of the terms used in the so-called descriptive sciences like astronomy and statistical thermodynamics and even our daily conversation are really theoretical terms, developed through a process of synthesis, and which describe objects which no one has even "seen." For example, a three dimensional cone is seen as a triangle, circle, ellipse, etc., from different angles, but never as a cone. The spherical shape of the sun is inferred on the basis of a number of lines of evidence, such as its circular shape, the flattened appearance of sunspots near the sun's edges and many others. The unobserved temperature at the center of the earth is handled in geology in much the same way as the unobserved ice ages.

But this is not the point of the paper. Assuming that the dichotomy can be justified, still, in the actual practice of science there has been no satisfactory way of stating in advance what kind of explanation would prove most fruitful, and therefore in every science both a descriptive approach and a hypothetical approach have been used, either successively or concurrently.

Einstein, for example, in his theory of relativity, presents what is essentially a descriptive theory in that underlying entities are not assumed. Yet this same scientist used a hypothetical and atomistic approach in his studies of Brownian movement, of the photoelectric effect, and the quantum theory of specific heats.

This may also be illustrated by the development of atomic theories during this century. Bohr's theory of a planetary atom, with certain assumed quantum properties, was an hypothetical approach which was highly successful in accounting for the properties of the hydrogen spectrum but which ran into insuperable difficulties in the attempt to apply the model to heavier elements. Born and Heisenberg, in their "matrix mechanics," then adopted a strictly phenomenological view in which only the observed characteristics of the spectral lines such as their intensities and polarization were manipulated mathematically, and this model was more successful than the earlier one. The development of Schrodinger's wave mechanics was a return to the hidden mechanism view, in which the physicists based their deductions on assumed properties of unobserved "things."

These periodic swings are part of the history of science and there is no reason to think that they will be different in the future. Both kinds of approaches are really supplementary, for there is no way of knowing whether we have tapped nature's hidden mechanisms until we try to follow out the implications of such assumptions. Some of the unobserved entities of physics have proved unfruitful and have been given up; the ether was such an entity as was an assumed "magneton" or atom of magnetism which several physicists postulated in an attempt to account for certain spectral phenomena. No empirical experiments were ever able to detect either the ether or the magneton and other explanations were eventually adopted. This was not the case for other assumed entities. The charge on the electron, for example, has been determined by at least 11 different methods, all yielding the same value. Avogadro's number (the number of molecules in a specified volume of gas at standard temperature and pressure) has been determined in many different ways (studies of Brownian movement, of the blue color of the sky, of the specific heat of solids, of equilibrium radiation, and so on) and always with identical results.

In fact, the essential criterion of the adequacy of a hypothetical construct is whether or not its assumed properties can be measured and checked in several different ways. This also is the basic difference between the intervening variable and the hypothetical construct.

This difference is something which can only be determined by experiment. In other words, only the multiplication of corroborative evidence will decide whether a scientist may call his terms hypothetical constructs.

Of course it is possible for one hypothetical construct to be substituted for another. The epicyclical paths of the planets, whose diameters and positions were computed by the Greeks, gave way in time to the simpler and more comprehensive hypothetical construct which assumed circular paths for the planets around the sun. In both cases the paths themselves are unobserved and their properties, such as their diameter, precession, etc., can be determined by several different methods.

In psychology there has been a development parallel to the other sciences. In the field of learning for example, many of the writers of the last two decades tried to relate each empirical finding to an assumed neurological model. A common construct was the "decrease in synaptic resistance" which was supposed to account for changes due to practice. To test this construct would require a definition of the term synaptic resistance and a method of measuring this independently of the performance changes being called "learning." This was never specified. Among the more recent mathematical learning theories, that of Estes (1), for example, discusses certain kinds of inferred stimulus-response bonds which relate an operationally unmeasurable stimulus element to an unidentifiable unitary response. Though the equations fit some data, the underlying constructs may have nothing to do with "reality" in the sense that there could be several ways of measuring the properties of the assumed connection.

On the other hand, the work of Tolman (11) was a relatively early attempt to get back to what is largely a descriptive psychology of learning. Tolman generally avoided neurological notions, and the theoretical terms which he did use were based upon the "facts and laws of learning" and an analysis of "molar" behavior. The work of Skinner (9) is well known as a more recent abstractive approach to the same problems.

In the area of perception a brain model has generally been invoked as the basis for an explanation of observed relations, for example, the Gestalt theory of isomorphism. The recent book by Gibson (2) however, is a return to a descriptive approach. In it he tries to develop a "psychophysical" theory of perception based on the relations between the responses of an observer and the characteristics of physical stimulation. This is an attempt to arrive at lawfulness through a precise description of what we see, without reference to any neurological or physiological processes.

This two-fold approach is pertinent to many of the concepts used in psychology: some psychologists for example, have sought the biological basis of such terms as intelligence, memory, and libido; others have conceived of these same terms as simply summary words for certain observations, the words not having any reference to underlying entities.

Generally, however, the hypothetical constructs used in psychology have never been very adequately formulated. Psychologists also have shown little interest in testing and determining the implied constants underlying behavior. Neither Hull nor his students, for example, have shown much concern with determining the characteristics of "neural oscillation" or of the "excitatory potential" which seem so relevant to his theory.

The Gestalt psychologists likewise, with some notable recent exceptions, have seldom done anything about directly studying the implied neurological correlates of the behavioral field.

In a sense this whole concern with intervening variables and hypothetical constructs, with descriptive science and causal science, is part of the more general problem of the "reality" of what is implied by our theoretical concepts. A term like "drive" may be a convenient summary word for certain operations of deprivation as related to certain kinds of responding or it may refer to certain changed conditions in the body; but if used in the latter sense, there are a number of implications which might be independently checked. It should then be possible to determine and produce the state of drive in more than one way, and one of the constants characterizing an individual should thereby be open to statement.

In summary, it should be clear that the question is not one of hypothetical constructs *or* intervening variables; let each theorist or experimenter use what mode of explanation seems to fit the facts best or is most useful in suggesting further experiments, and let our continuous questioning of nature eventually decide whether we have come upon the hidden mechanisms underlying our phenomenal world, or whether we can get the clearest and most encompassing picture of our universe by sticking to what we see.

## REFERENCES

1. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94-107.
2. GIBSON, J. J. The perception of the visual world. New York: Houghton Mifflin, 1950.
3. HODGES, H. A. Wilhelm Dilthey, an introduction. London: Kegan Paul, 1944.
4. KESSEN, W., & KIMBLE, G. A. Dynamic systems and theory construction. *Psychol. Rev.*, 1952, **59**, 263-269.
5. KRECH, D. Dynamic systems, psychological fields, and hypothetical constructs. *Psychol. Rev.*, 1950, **57**, 783-790.
6. MACCORQUODALE, K., & MEEHL, P. E. On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.*, 1948, **55**, 95-107.
7. MEYERSON, E. Identity and Reality. New York: Macmillan, 1930.
8. RANKINE, W. J. M. Outlines of the Science of Energetics. (Ch. 12 in *Miscellaneous Scientific Papers.*) London: Griffin, 1881.
9. SKINNER, B. F. The Behavior of Organisms. New York: Appleton-Century, 1938.
10. ————. Are theories of learning necessary? *Psychol. Rev.*, 1950, **57**, 193-216.
11. TOLMAN, E. C. Purposive Behavior in Animals and Men. New York: Appleton-Century, 1938.
12. ————. Symposium on perception and personality. *J. Personal.*, 1949, **18**, 48-50.

*Hofstra College*
*Hempstead, New York*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 65-73.

# SENSORY VS. AUTONOMOUS CONTROL OF SPAN OF APPREHENSION*[1]

*Payne Whitney Psychiatric Clinic, New York Hospital*

---

CHARLES A. KNEHR AND NUNI FULLER

---

## A. THE PROBLEM

The development of individual cerebral organization as a result of experiences from infancy is the basis for psychological functioning of the human organism. The patterning and repatterning of neuronal circuits appears to be in terms of the fortuitous events of experience compounded with and partially determined by the organization of the moment. The later yield in a healthy mature individual is an astounding variety of behaviors. Some patterns become stable enough that the individual comes to present many predictable characteristics which differ and set him off from other individuals. Even with the appearance of stability these behaviors are not ordinarily fixed and machine-like, but are responses that occur selectively to stimulation in terms of patterns of varying alterability now built into the organism itself. Names used to denote these central effects are set, expectancy, constancy and the like (2).

It has been clearly indicated from William James onward that the lack of one to one relationship between sensory stimulation and perception or the lack of fixity between stimulus and response provides the psychologist with his essential phenomena for study. A response must be a resultant of factors including the sensory input and the cerebral activity of the moment. The latter always includes the lasting alterations in cerebral organization resulting from previous interactions of the same kind and is referred to as learning.

It has been suggested that cerebral events may possibly escape from sensory control so that the active factors in the resultant perception or behavior would be central nervous events (2). It is difficult to conceive of this situation except as a matter of degree. The extreme might be represented by autonomous events relatively free of sensory control. Such a situation might obtain for only brief periods, if at all, in a mature healthy individual. However, it would seem possible that the pathological conditions represented by dissociative

---

*Received in the Editorial Office on July 17, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] From the New York Hospital and the Department of Psychiatry, Cornell University Medical College, and Hunter College of the City of New York.

phenomena and schizophrenic characteristics might represent this extreme when held for extended periods.

There appear to be several ways of attacking the problem of the relative strength of sensory or central control in behavior or perception. Our immediate interest in the problem is from the general point of view of the study of sensitivity in individuals to stimulus conditions which could provide a set which in turn would affect performance or perception. We have a long term interest in the quantitative study of the relation between sensory and central factors in different psychopathological processes involving thinking disorders and deterioration.

The present study attempted to answer two questions. First, what effect would varying the size of the area over which the stimuli appear have on the response to groupings of dots presented by means of a short exposure device? And secondly, would there be some differential effect in the responses to the stimulus variations used between patients with varying psychopathology and persons without obvious pathology? In general, we were trying to see how sensitive individuals would be to the compression of stimulus material into different sized fields when no specific instruction in reference to field size was given.

## B. Procedure

The apparatus consisted of a film strip projector made by the Spencer Lens Company to which an Ilex shutter was attached. Projection was from behind and through a translucent screen, 6x8 inches in size, set in a wall separating the subjects from the apparatus. The exposure was 20 milliseconds. The dots, when projected were white on a dark field, measured 4 mm. in diameter and subtended an angle of about 0.3° at the eye from a viewing distance of 75 cm. The luminance of the dot figures was approximately 14 foot-lamberts, with a field of about 0.5 foot-lambert from the viewing side of the screen.

The stimulus material was recorded on a film strip of 48 frames made in the following manner. A series of dot patterns was prepared on drawing paper, there being six each of three to 10 dots inclusive. The six patterns for each number of dots were further subdivided so that two of the dot patterns were scattered so as to occupy the entire viewing area, two in 42 per cent of the central portion of the viewing area, and two in which the dots were compressed into 14 per cent of the area. The proportions of the screen were maintained for each field size. The placement of dots within the appropriate field was random. This was achieved by allowing small cardboard bits to flutter down onto a template. When the required number about filled the allotted space, the positions were marked and inked in. The 48 patterns made one by one in this manner were then randomized for order of presentation and photographed. The serial order was the same for every test.

On being positioned before the screen, the following written instructions were presented for the subject to read:

> On the screen in front of you, pictures of dots will be exposed for brief intervals.
>
> At the signal "Ready" you will look at the small cross in the center of the screen. A few seconds later the dots will be briefly exposed on the screen.
>
> All you have to do is report the number of dots seen. If you are not sure, guess and give a single number.

Thirty-nine subjects were tested. Twenty control subjects were students at Hunter College and Cornell University Medical College. Nineteen were inpatients at the Payne Whitney Psychiatric Clinic with varying psychopathology differing in severity. Only patients capable of coöperation were tested. All subjects answered the Minnesota Multiphasic Personality Inventory (MMPI), mainly as a check on the control subjects. Psychopathology of patients at the time of testing was evaluated clinically. To estimate the consistency of the obtained results, five controls and five patients were retested two to 15 weeks following the original test.

Where the experimental subjects were grouped for purposes of statistical comparison, the controls were taken all together ($n=20$), patients with schizophrenic characteristics as the main feature were grouped ($n=9$), while the remaining patients with anxiety as an essential feature were grouped together ($n=10$). In order to assess possible relationships between individual characteristics of responses and patterns of pathology, an index was prepared for each control subject and each patient. Linearity was assumed for the data. The index was the slope of the line of best fit multiplied by the scatter around that line (standard error of estimate computed from $r$).

## C. RESULTS

The effect of the different field sizes on the subjects' estimates of the number of stimulus dots in each pattern is shown by the data of Table 1. The largest increment of the medium field size over the small was 5.1 per cent. The increments of the large field over the small varied from 2.5 to 4.6 per cent for the several groups. It is seen that there is some overall tendency for the medium and large fields to increase the mean of the responses but the increments are obviously small. $T$-ratios were mostly near 1.0 with the largest being 1.56 for the increment between the small and medium fields for the control group. Study of responses by individuals would indicate no consistent trend related to field size. This is borne out by the low *t*-ratios. The results provide no evidence to indicate that the spatial spreading of stimulus dots has any appreciable effect per se on responses to tachistoscopically presented fields of dots.

TABLE 1
GROUP RESPONSE MEANS TO STIMULUS DOT PATTERNS IN SMALL, MEDIUM, AND LARGE FIELDS

| Number of dots in | Controls | | | Anxiety patients | | | Schizophrenic patients | | |
|---|---|---|---|---|---|---|---|---|---|
| stimulus pattern | Small | Medium | Large | Small | Medium | Large | Small | Medium | Large |
| 3 | 3.13 | 3.08 | 3.13 | 3.15 | 3.05 | 3.20 | 3.11 | 3.22 | 3.22 |
| 4 | 4.23 | 4.15 | 4.25 | 4.20 | 4.10 | 4.45 | 4.39 | 4.56 | 4.50 |
| 5 | 5.18 | 5.63 | 5.05 | 5.15 | 5.40 | 4.75 | 5.56 | 5.67 | 5.11 |
| 6 | 6.28 | 6.40 | 6.25 | 6.05 | 6.05 | 6.25 | 6.06 | 6.56 | 6.94 |
| 7 | 7.43 | 7.05 | 7.43 | 6.90 | 7.10 | 7.10 | 7.39 | 7.89 | 7.78 |
| 8 | 7.75 | 8.33 | 7.98 | 7.30 | 7.70 | 7.80 | 8.00 | 7.61 | 8.44 |
| 9 | 8.08 | 9.20 | 8.78 | 7.90 | 8.60 | 8.45 | 8.39 | 9.00 | 8.83 |
| 10 | 9.23 | 9.80 | 9.70 | 8.85 | 9.25 | 9.35 | 8.94 | 9.94 | 9.44 |
| Δ | | 4.5% | 2.5% | | 3.6% | 4.6% | | 5.1% | 4.6% |

*Note*: Largest *t*-ratio was 1.56 between small and medium fields for control group.

In the course of analyzing the data on individuals, it was found that different subjects, both controls and patients alike, showed what appeared to be highly individual modes of response when their responses were plotted against the correct stimulus values. Examples of two quite different response scattergrams are shown in Figure 1. To check the stability of these individualized responses, 10 subjects (five controls and five patients) were retested and the new data plotted. Direct comparisons of such retests made from two to 15 weeks after the original test showed a remarkable similarity in pattern between the two sets of response data on the same individuals (Figure 2).

The finding that patterns of responses had an individual stamp which appeared to hold up on retest, made it seem reasonable that the individual character of the response patterns might relate to some one or several factors in personality makeup or perhaps to some psychopathological characteristic. The comparisons seemed to show that patients with schizophrenic illnesses gave more scattered responses and higher slopes than other subjects in spite of a number of negative instances. More precise evaluation of these data were needed to show whether this was anything more than an impression.

When the indices (see above) for the individual scattergrams were combined into three groups, the means were: control group, .78, schizophrenics .94, and anxiety group .62. The largest *t*-ratio was 1.86 between the latter two and fell short of significance. It was apparent from the individual indices that there was a good deal of individual variability within each of the groups and that individual response patterns of similar nature would be found among all subjects no matter what their diagnostic labels.

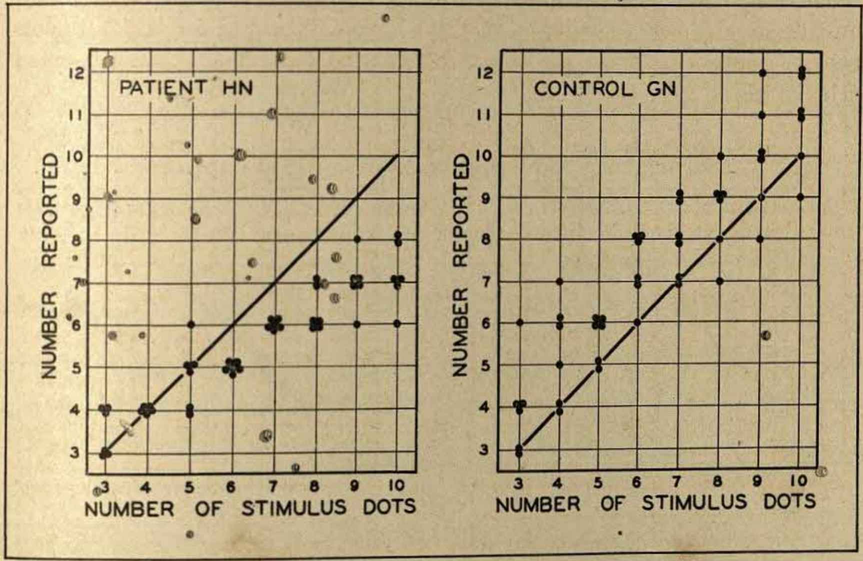


FIGURE 1

Sample Response Scattergrams to Random Patterns of Dots Presented Tachistoscopically. The data include responses for all three field sizes. Correct responses would lie along the diagonal line.

Direct comparisons of response scattergrams for individual patients showed no observable trends. The negative instances inevitably appeared to balance the positive. Quantification of these same data made it possible to correlate characteristics of the response pattern with *t*-score values on the MMPI.

Inspection showed little or no correlation between individual indices and separate scales on the MMPI. Wherever there appeared to be some possibility of a relationship Pearson *r*'s were calculated and found to be very small and not significantly different from zero. We did obtain a slightly higher *r* when the indices were set against a combination of three MMPI scales (*F*, *Pt*, and *Sc*) which in our experience are often elevated together in patients with schizophrenic illnesses. When combined arithmetically, the correlation with individual indices was $r=.36$. The standard error was so large ($\sigma_r=.21$) that we cannot assume a relationship on the basis of the data shown. Individual response patterns were not clearly related to other factors evaluated in this study.

## D. Discussion

So far as the authors have been able to determine, relatively few recent experiments have been made with the span of visual apprehension method dating back well beyond the turn of the century (1, 6, 7). We started with

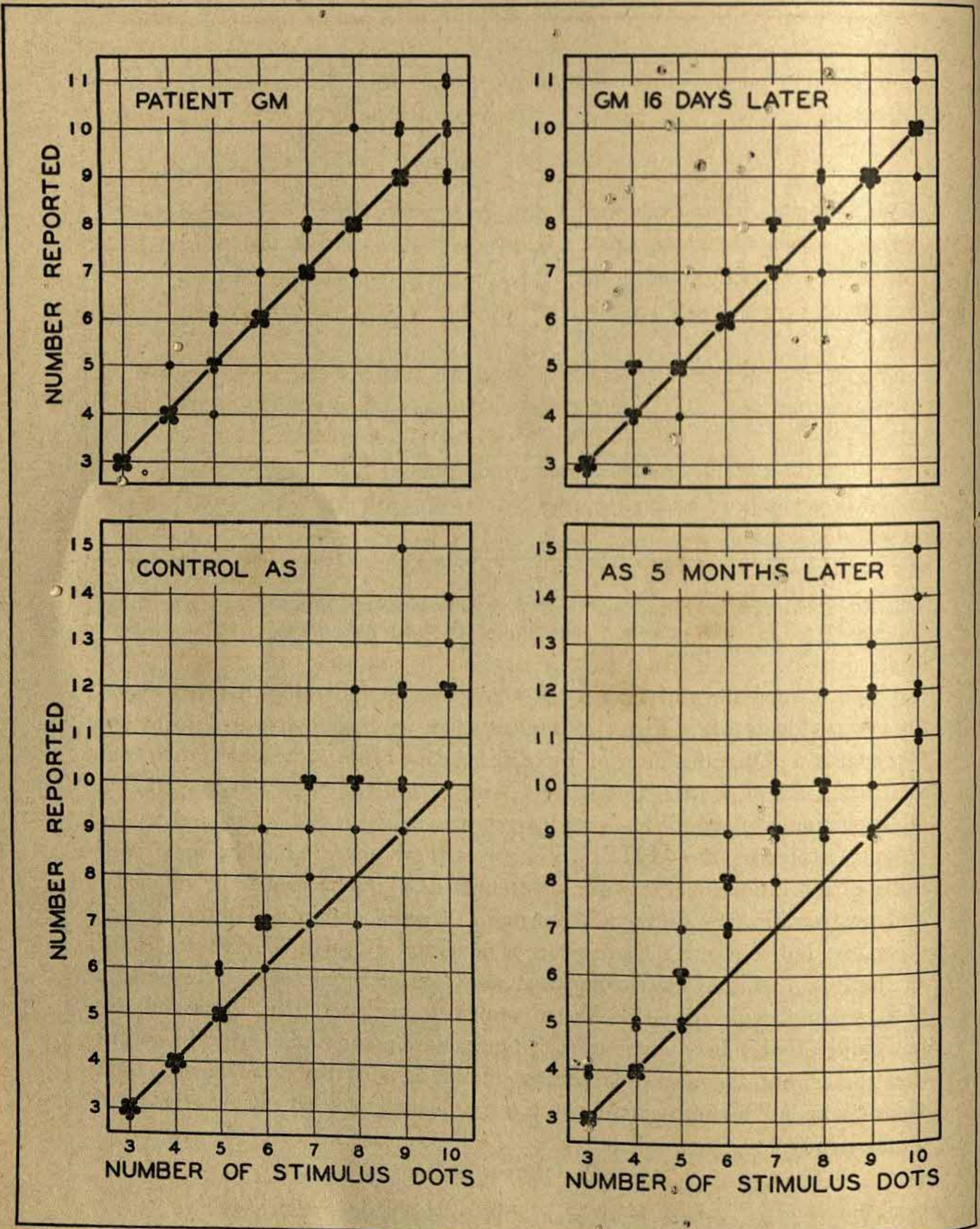


FIGURE 2
The Persistent Individual Character of Response Scattergrams Shown on Retest

the assumption that the technique would provide us with a means of measuring the development of uninstructed sets related to the size of the area over which stimulus dots were visible. Should such sets occur differentially from one individual to another, relationship to certain psychopathological characteristics might be found.

When statistically treated, our data do not support the first hypothesis. Slight increases in the number of dots reported for the larger fields (in the direction of the true stimulus values) did not approach significance. The accuracy of reporting the number of stimuli shown on short exposure is in general agreement with the work of previous investigators where the method is similar.

While individual differences have been the rule in earlier experiments, the interesting finding of the present investigation is the individual patterning of responses seen when the subject's estimates are plotted against the correct stimulus values. The resulting scattergram varied from person to person but in the same individual was similar on retest. While in the course of the present experiment we have not found variations in individualized patterns of responses that are clearly related to personality characteristics, further work may produce related factors. Persistent patterns suggest that central factors determine to an unknown but marked extent the way in which the stimulus patterns are responded to. Some individuals are capable of correct responses throughout the range of three to 10 stimulus dots with relatively few errors. Others tend to overestimate the smaller values and underestimate the larger. Occasional individuals consistently make errors of overestimation to stimuli of all sizes with a relatively high degree of variability. Such persistent characteristics must ultimately be accounted for.

Because of the differences in experimental procedure, it is difficult to relate these results to the recent work of Hunter and Sigler (4). A little speculation, however, may not be out of order. These experimenters have explained their data on perceptual discrimination in terms of time and intensity factors on the receptor side. We have found individual differences in response to a presumably constant situation which suggest, on the surface at least, that span is a function of factors that will include intensity and time relationships but that these latter are not sufficiently determining alone. Conceivably, central facilitation or interference and possible overloading of pathways leading from visual cortex to areas beyond are involved. On the surface, this suggestion would be somewhat akin to Schlosberg's discussion of the Hunter-Sigler experiment (5). The differences between individuals could represent differences in central functioning as readily as those in receptor function. Persistent and similar individual characteristics in the two subjects studied by Hunter and Sigler might not yield response data suggestive of the possible rôle of central factors in span of apprehension experiments. Too much organism seems to be interposed between stimulus and response for receptor function to be so thoroughly controlling in a situation involving perception of and response to complex stimulus patterns.

It is this very problem, i.e., the relative influence of sensory and autonomous factors in behavior that ultimately may be of some importance in gaining an understanding of certain perceptual and thinking changes in clinical patients. Such relationships were implied for example, in Hollingworth's application of the term "sagacity" (3). Dissociative phenomena and the "split" in schizophrenia imply differential characteristics which may represent deviations of sensory-autonomous balance from some unknown optimal level. It is possible that retinal information is carried quite accurately as far as the projection areas. From that point on, individual cerebral organization will determine in an integrative way what can lead off to consistent response or how much data not specifically patterned can be handled before mutual interference of greater or lesser extent results. Already active neurone networks could positively or negatively facilitate information from the receptor and determine to a large extent in this experiment whether an accurate response would or would not be made. In the present experiment presumably we have delivered visual data of increasing complexity to the visual cortex under the conditions of a span of apprehension experiment. We are impressed with the individual character of the patterns of response and the need for explanation even though we have been unable to relate the persistent patterns to any measured or observed variable.

## E. SUMMARY

A span of apprehension experiment using 39 subjects was designed to test the hypothesis that varying the size of the visual area over which stimulus dots appeared would alter the response to those stimuli. A second hypothesis assumed that the effect would be differential between control subjects and persons with schizophrenic or other psychopathological characteristics.

The evidence provides no significant support for either hypothesis.

On the positive side, we have found indications of persistent individual patterns of responses when the results were plotted as scattergrams. The patterns remained on retest from two to 15 weeks later in 10 subjects. The factors underlying these individual patterns are unknown to us. No significant relationships were observed between the individual response characteristics and available personality data.

## REFERENCES

1. GRAHAM, C. H. Visual perception. In S. S. Stevens (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951. (Pp. 868-920.)
2. HEBB, D. O. The organization of Behavior: A Neuropsychological Theory. New York: Wiley, 1949.
3. HOLLINGWORTH, H. L. Abnormal Psychology. New York: Ronald Press, 1930.
4. HUNTER, W. S., & SIGLER, M. The span of visual discrimination as a function of time and intensity of stimulation. *J. Exp. Psychol.*, 1940, **26**, 160-179.
5. SCHLOSBERG, H. A probability formulation of the Hunter-Sigler effect. *J. Exp. Psychol.*, 1948, **38**, 155-167.
6. WHIPPLE, G. R. Manual of Mental and Physical Tests. Part I: Simpler processes. Baltimore: Warwick & York, 1924.
7. WOODWORTH, R. S. Experimental Psychology. New York: Holt, 1938.

*Payne Whitney Psychiatric Clinic*
*New York Hospital*
*525 East 68 Street*
*New York City 21*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 75-80.

# THE INFLUENCE OF THREAT APPEALS ON SELECTIVE LEARNING OF THE CONTENT OF A PERSUASIVE COMMUNICATION*[1]

*Department of Psychology, Yale University*

IRVING L. JANIS AND HARRY C. MILHOLLAND, JR.

## A. THE PROBLEM

Among the various types of contents used as motivating appeals in persuasive communications are those which emphasize potential dangers or deprivations that are alleged to result from failure to adhere to the communicator's recommendations. Such contents, which will be referred to as "threat appeals," are likely to stimulate at least a mild degree of emotional tension and may sometimes elicit intense fear.

Observations of reactions to various types of fear or anxiety stimuli suggest that when emotional tension is strongly aroused, attention, concentration, and learning tend to be temporarily impaired [cf. Hanfmann (1)]. If this were to occur when a threat appeal is presented in a persuasive communication, the audience would be less likely to learn the message and consequently would be less influenced by it.

Potential inferences of this kind were investigated by Janis and Feshbach (2) in an experiment which compared the reactions of equivalent groups exposed to strong, moderate, and minimal threat appeals in a dental hygiene communication. The main findings indicated that the use of a strong appeal lowered the effectiveness of the communication, producing less opinion change than the milder appeals. On the basis of supplementary observations, the authors inferred that the interference was primarily a matter of motivational changes rather than a learning decrement: as a consequence of having been left in a state of partially unrelieved emotional tension, the group exposed to the strong appeal apparently became motivated to ignore or to minimize the importance of the threat. The degree of interference attributable to inattentiveness or to other sources of reduced learning efficiency appeared to be negligible inasmuch as all experimental groups learned the factual material on

---

*Received in the Editorial Office on July 25, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was conducted at Yale University, as part of a coördinated program of research on communication and opinion change. The Yale communication research project is under the general direction of Professor Carl I. Hovland and is financed by a grant from the Rockefeller Foundation.

dental hygiene equally well, as measured by a comprehensive information test given immediately after exposure to the communication.

In the present experiment an attempt was made to investigate further, with a more sensitive measure, the effects of threat appeals on attention and learning. It was felt that by using a verbatim recall test (rather than an objective test which mainly taps recognition memory) it might be possible to detect more subtle decrements. The purpose was not only to check on the earlier findings regarding overall learning efficiency but also to compare strong and mild appeals with respect to selective effects on the type of content that is learned.

## B. Method

The experiment was designed to compare the verbatim recall of subjects exposed to two forms of the same persuasive communication. Form *A*, which presented a great deal of threat material, was 270 words in length; while Form *B*, which omitted much of the threat material, contained 190 words. Both forms were based on excerpts from the illustrated talks on dental hygiene which had been used in the earlier experiment (2) and were presented as a short magazine article from an authoritative journal.[2] They contained the same three major themes:

1. An introductory statement which asserted that many people do not gain the full benefits of toothbrushing because they brush their teeth at the wrong time.
2. A central argument to the effect that cavities and other damage to the teeth are caused by the food particles which remain after eating.
3. A concluding recommendation that one should brush his teeth immediately after each meal.

In addition, both forms presented seven supplementary assertions, one of which was a statement of the original source of the material presented (the *Journal* of the American Dental Association). Three of the supplementary assertions provided an elementary explanation of the reasons why food particles cause damage (they foster the growth of bacteria; they form strong acids; the acids dissolve the enamel coat of the teeth). Three other assertions referred to specific threatening consequences of failing to adhere to the recommendation (sore or inflamed gums, mouth discomfort, need to have cavities treated by a dentist).

In both forms, the 10 content items were worded in the same way. Form *A* (the strong appeal) differed from Form *B* (the mild appeal) only in that it contained material which elaborated on the dangerous consequences of developing unhealthy teeth: violent toothaches; painful sessions of drilling; dental surgery in order to cut away completely decayed teeth; suffering from bleeding gums; possibility of developing serious mouth infections which could result in severe illness, prolonged fever, and loss of all the teeth.

---

[2] Copies of the two forms of the communication are available from the authors upon request.

Form *A* was given to a well-educated group of 24 adults, including professional men and women, middle class housewives, and college students; Form *B* was given to an equivalent group of 22 subjects. The two groups were carefully equated with respect to age, sex, educational level, and occupation.

Identical procedures were used with both groups. Every subject was seen in an individual session and was asked to read through the communication once, in the way he would ordinarily read any short article in a magazine. (Although no time limit was imposed, there was relatively little variation in reading time from one subject to another and the mean reading time of the two groups was found to be approximately the same.) After the subject had finished, the article was removed and he was told that, although he had not been asked to memorize it, the experimenter was interested in finding out how much of the material could be remembered. The subject was then asked to write down the passage as completely as possible and to make an effort to reproduce word for word as much of it as he could.

The papers were scored according to whether or not each of the 10 content items was correctly recalled. The subject was credited with correct recall of a given item if he stated the gist of the assertion, irrespective of whether or not he reproduced the original wording. In order to check on the reliability of the scoring procedure, all papers were scored independently by two judges. Out of a total of 460 judgments (46 subjects × 10 content items) there was agreement on 434 (i.e., 94.4 per cent agreement). The relatively small number of disagreements did not cluster within any given content category. They were resolved by formulating additional scoring rules, without taking account of whether the subject had been exposed to Form *A* or *B*.

## C. RESULTS

Results bearing on the overall performance of the two groups are presented in Table 1. It is apparent that the mean number of content items correctly reproduced is almost identical in the two groups and that approximately equal percentages stated correctly each of the three major content themes. Since the differences on these measures are negligible, it appears that the inclusion of the additional threat material in Form *A* had no general effect on attention or learning of the main content of the communication.

TABLE 1
Overall Recall Performance of Group $A$ and Group $B$

| | Group $A$: Exposed to strong threat appeal ($N=24$) | Group $B$: Exposed to mild threat appeal ($N=22$) |
|---|---|---|
| 1. Mean number of content items correctly recalled | 6.0 | 5.9 |
| 2. Percentage who correctly recalled the introductory theme: Many people do not brush teeth at right time | 79% | 77% |
| 3. Percentage who correctly recalled the main argument: Remaining food particles cause damage to teeth | 96% | 100% |
| 4. Percentage who correctly recalled the main recommendation: Brush teeth immediately after each meal | 100% | 100% |

TABLE 2
Percentages of Group $A$ and Group $B$ Who Correctly Recalled Each of Seven Supplementary Content Items

| | Content items | Group $A$: Exposed to strong threat appeal ($N=24$) | Group $B$: Exposed to mild threat appeal ($N=22$) | Difference ($A$—$B$) | p-value (two-tail test) |
|---|---|---|---|---|---|
| Threatening consequences | 1. Sore or inflamed gums | 75% | 59% | +16% | .25 |
| | 2. Mouth discomfort | 38% | 9% | +29% | .02 |
| | 3. Need to go to a dentist | 75% | 14% | +61% | $<.01$ |
| Explanatory assertions | 4. Remaining food particles foster growth of bacteria | 25% | 64% | —39% | $<.01$ |
| | 5. Remaining food particles form strong acids | 67% | 68% | — 1% | $>.90$ |
| | 6. Acids dissolve the enamel coat | 37% | 68% | —31% | .04 |
| Source of the communication | 7. *Journal* of the American Dental Association | 4% | 27% | —23% | .03 |

Analysis of the supplementary content items, however, revealed marked differences between the two groups. Table 2 shows the results for all seven supplementary items, grouped according to type of content. It will be noted that Group $A$ is superior to Group $B$ on the three items dealing with specific threatening consequences. In the case of two of the three items, the percentage difference is large enough to be statistically reliable. But on the remaining content items, the differences are in the reverse direction. On two of the explanatory assertions which dealt with the causes of the threat, Group $A$

shows reliably poorer recall than Group *B*. Similarly, a reliably smaller percentage of Group *A* correctly recalled the source of the communication. In general, the results in Table 2 indicate that the strong threat appeal, as compared with the mild one, had a selective effect upon the type of information that was acquired from the communication.

## D. DISCUSSION

The overall recall performance of the adult subjects in the present experiment (Table 1) shows the same outcome as the recognition test performance of the young adolescents observed in the earlier experiment on effects of fear-arousing communications (2). This replication of the earlier findings provides additional support for the conclusion that normal personalities can ordinarily tolerate threat appeals—of the sort which play up familiar dangers concerning pain and body damage—without manifesting any marked impairment in general learning efficiency.

Nevertheless, the recall data on the supplementary content items (Table 2) indicate that the use of a strong threat appeal does exert some influence on the learning of the communication. From the pattern of the results, we infer that the group exposed to the mild appeal was more likely to perceive the communication as a scientific report and to set themselves to retain the explanatory content and the source of the information. Among those exposed to the strong appeal, on the other hand, attention seems to have been focused primarily on the threatening consequences. The findings suggest that when a communication deals with relatively minor threats, as in the case of the dental hygiene material used in this study, elaborating on the potential dangers has the effect of making the threat content more vivid and therefore better learned; but the heightened tendency to learn the threat content tends to interfere with the learning of other material contained in the communication—in this case, material on the causes of the threat.

The selective learning tendency observed in this experiment could sometimes operate in the direction of augmenting the degree of attitude change, inasmuch as better learning of the threatening consequences might increase the audience's motivation to adhere to the communicator's recommendations. Under certain circumstances, however, the opposite outcome would be expected. Consider, for example, the use of threat appeals in persuasive communications that are designed to prepare people for coping with future danger situations. The audience's tendency to retain primarily the threat content highlighted by a strong appeal would be particularly detrimental if adherence to the recommended action is contingent upon comprehending and taking account of information about the causes of the threat.

Thus, one of the ways in which the use of strong threat appeals may affect attitude changes is by influencing what is learned and retained from the content of a persuasive communication. If the selective learning effects observed in the present experiments are verified, further research will be required to determine the conditions under which this effect facilitates or interferes with the ultimate effectiveness of persuasive communication.

## E. Summary and Conclusions

This experiment consisted of comparing the verbatim recall performances of two equated groups of adult subjects, one of which read a short magazine article on dental hygiene containing a strong threat appeal while the other read another version of the same communication containing a mild appeal. When tested immediately after exposure to the communication, the two groups did not differ in the mean number of content items correctly recalled nor in the accuracy of recall of each of the major content themes, including the recommended conclusion. These findings agree with those of an earlier experiment (2) in indicating that there is no marked loss in overall learning efficiency as a result of using strong threat appeals of the sort which emphasize familiar dangers regarding pain and body damage.

Additional findings from the present experiment, however, indicate that specific assertions about unfavorable consequences of the threat were recalled by a significantly higher percentage of the group exposed to the strong appeal; whereas, explanatory assertions about the causes of the threat and a statement which specified the source of the information were recalled by a significantly higher percentage of the group exposed to the mild appeal. These findings reveal a selective recall tendency and suggest that threat appeals can exert an important influence on what is learned and retained from the content of a persuasive communication. The implications of this selective tendency were discussed with regard to the effectiveness of communications that are intended to modify beliefs and attitudes.

## References

1. Hanfman, E. Psychological approaches to the study of anxiety. In P. H. Hoch and J. Zubin (*Eds.*), *Anxiety*. New York: Grune & Stratton, 1950. (pp. 51-69.)
2. Janis, I. L., & Feshbach, S. Effects of fear arousing communications. *J. Abn. & Soc. Psychol.*, 1953, **48**, 78-92.

*Department of Psychology*
*Yale University*
*333 Cedar Street*
*New Haven, Connecticut*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 81-93.

# DIRECTIONAL AFTER-EFFECTS FOLLOWING SYSTEMATIC DISTORTION OF THE VISUAL FIELD*

*Department of Psychology, University of Washington*

SIDNEY S. CULBERT[1]

## A. THE PROBLEM

It has long been recognized that many factors are involved in the determination of the principal axes of orientation in the visual field, the horizontal and the vertical. For convenience these factors are sometimes divided into "*g*-factors" and "*v*-factors," the first being all those resulting directly or indirectly from the pull of gravity, the second, all those arising from visual stimulation. Additional factors such as suggestion and expectation, which are not properly classifiable into the categories mentioned, are usually of minor importance.

Gibson and Mowrer have defended the viewpoint that the axes of orientation are determined by visual factors and gravitational factors acting jointly, but that the vertical and horizontal axes of perceived space have no empirical meaning except with reference ultimately to the direction of terrestrial gravity (2). On the other hand, Koffka has maintained that orientation depends upon the total organization of the visual field, that the main directions of the field, the vertical and the horizontal, constitute the framework (4, 215 ff.). Elsewhere he has held that those parts of the field which assume the function of framework will be seen as horizontal or vertical whether their retinal images are or not (5). Thouless agrees with Gibson and Mowrer that the two kinds of factors act together jointly, but gives neither group priority over the other. He considers the situation to be an example of a more general principle which he calls the law of phenomenal compromise—that when two sets of sensory or perceptual cues which alone would give rise to phenomenal characters inconsistent with one another are presented together, the phenomenal character which is actually experienced

---

*Received in the Editorial Office on August 1, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This experiment was a portion of a study submitted to the Department of Psychology, University of Washington, in partial fulfillment of the requirements for the Ph.D. degree. The writer is indebted to Dr. Erwin A. Esper for his general supervision of the study, and to Dr. Roger Loucks for suggestions on procedure and apparatus.

is neither that indicated by one nor that indicated by the other set of cues but is a compromise between them (7).

No one of these viewpoints can be true of all perceptual situations. The visual vertical can be established with great accuracy in the absence of visual cues, as was shown by Neal (6). On the other hand, *g*-factors may be entirely disregarded in favor of visual cues. If the visual field is the complex one of everyday life there are many objects the main lines of which establish the horizontal and vertical no matter what angle these make with the meridians of the retinal field. In this situation, however, the non-correspondence may be accompanied, as Gibson and Mowrer claim, by concomitant alterations of postural stimuli which could, so to speak, certify the continued acceptability of the visual framework despite its non-correspondence with retinal meridians. But in some cases the visual field establishes its framework without regard to retinal meridians and with no concomitant alteration of postural or other *g*-factor stimuli. The "haunted swing" illusion indicates the strength of the rotating visual field in determining *S*'s perceptual framework. Even knowledge of the real situation will not serve to dispel the illusion as Wood has pointed out (10). The *v*-factors show no compromise with *g*-factors in this example.

Many experimental attempts have been made to determine whether a visual field optically tilted by a fixed amount would be "accepted" by a subject even though not in accord with *g*-factor cues. Negative results have usually been reported, though it is not always clear what criteria were being used by *O*s to determine acceptability. Apparently Wertheimer's tilted-mirror field established acceptable phenomenal horizontal and vertical for his subject in a short time (9), but others who have tried the experiment do not agree—as Gibson and Mowrer suggest, ". . . possibly Wertheimer was reporting a relatively more cognitive phenomenon—the decrease in the strangeness and unreality of the mirror-scene" (2, p. 310). In other experiments where prism lenses rather than a mirror were used to tilt the visual field *S*s became able, after a period of adjustment, to take part in ordinary activities nearly as well as before, but none reached the point where the tilted field was felt to be phenomenally upright as had been the "real" world of his former experience.

In these experiments only one sort of transformation of the visual field was attempted—the tilting of the field around the visual axis, a kind of change which occurs in daily life with every tilt of the head. The experiments differed from normal visual experience in only three significant respects: there was no concomitant alteration of postural stimuli, the degree of tilt was greater and more persistent than commonly occurs in non-experimental situations, and *S* was more aware of his sensations. In some experiments these differences were, of course, very great. An inversion of the visual field (i.e., a "tilt" of 180°) is very different from the degrees of tilt normally experienced. But the visual field itself was not changed except in relation to *O*.

Even if the main lines of the tilted fields in these experiments had, in time, been completely accepted as the phenomenal horizontal and vertical of the subjects it could be contended that this showed only that visual factors are, in such cases, able to determine the framework regardless of inconsistent postural cues; it would show that the visual field can determine the phenomenal horizontal and vertical when its relationship to *O* is altered, but the question would be left open—if the visual field itself is distorted by a systematic transformation of its coördinates in a way which does not occur in any ordinary situation, will the main lines of the field influence *S*'s phenomenal horizontal and vertical? An affirmative answer to this question will lead to other problems the investigation of which should have certain implications for theories of perception. The three main questions which this study is designed to answer may be stated as follows:

1. If an *S* is placed in a situation in which *g*-factors are held approximately constant and is presented with a visual field, all figured parts of which are systematically distorted (by transformation) in a way which would not occur in natural surroundings, will *S*'s phenomenal horizontal be influenced in the direction of the coördinates of that system?
2. Will the effects, if they occur, persist after the distorted visual field and its *retinal* after-effects have entirely disappeared?
3. Will the effects, if they occur, be related to the degree of distortion of the visual field?

To investigate these questions an experiment was carried out, the purpose of which was to compare the directional after-effects of six geometric transformations of a rectangular grid composed of 10 vertical lines and 10 horizontal lines, the grid being the only figure in an otherwise black visual field.

## B. Apparatus and Procedure

Basically, the apparatus was a device for presenting to *S* a series of straight lines of different degrees of slant (test object) alternating with a visual design (induction design) which remained the same throughout a given series. The experimental room was light-proof, and the induction design and test object were white figures on black backgrounds to reduce reflected light to the absolute minimum. In addition, light traps and light shields faced with black velour paper were used to catch stray reflected light from the figures. In order to rule out the possibility that visual after-images of the induction design were present when the test object was shown, the former was presented to the right eye only, the latter to the left eye only.

All time intervals, signals, and variations in position of the test object were controlled automatically, although also subject to manual control during the instruction period. The plan of the apparatus is shown in Figure 1.

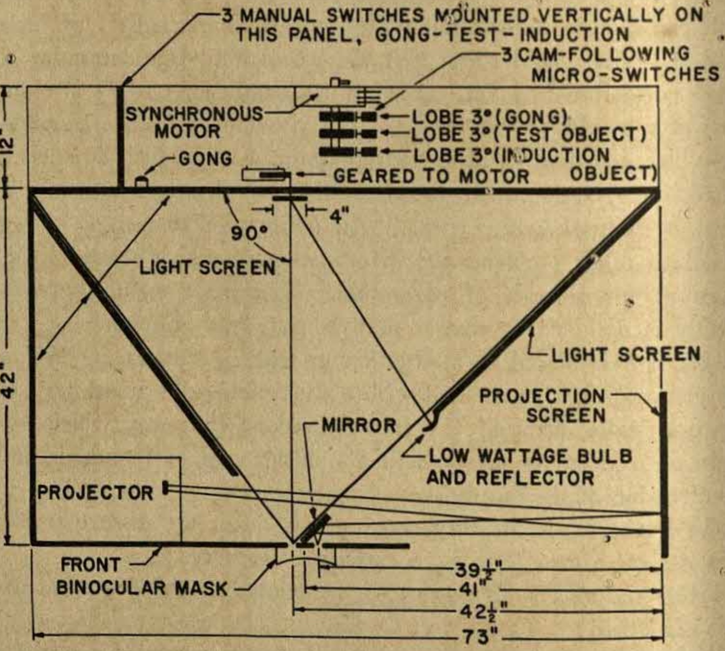


FIGURE 1
PLAN OF THE APPARATUS

The test object was a rectangular steel bar ½ by 4 in. painted white on the side visible to *S*. Its midpoint lay in the line of regard of the left-eye primary position, 42 in. from the eye. Each time the test object was presented to *S* by illumination from a source invisible to *S* its position relative to the horizontal was different. A 40-lobe cam guiding a follower attached to the shaft of the test object was synchronized with the cam controlling the source of illumination in such a way that the test object was stationary while visible, but moved to the next position during the period of darkness.

The 40 positions of the test object included every position from +1½° through horizontal to —5° in 10′ steps. Thus *S* responding in terms of the objective situation should see nine positions of the test object as slanting up to the right, one as horizontal, and 30 as slanting down to the right. The induction designs were six in number, one being the undistorted grid and the other five being five transformations of the grid (Figure 2). The trans-

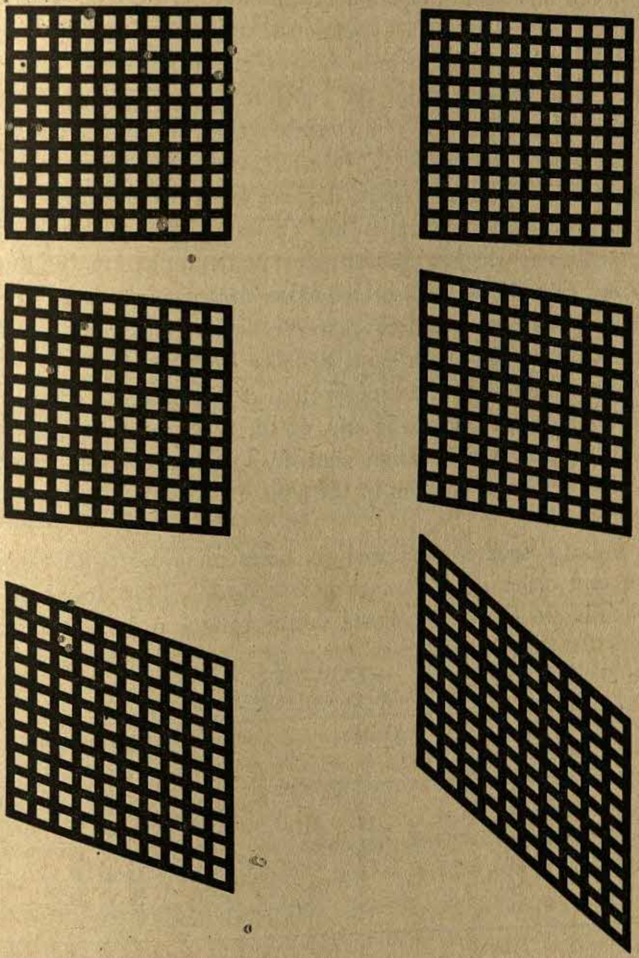


FIGURE 2

THE SIX INDUCTION DESIGNS PRESENTED TO *S*S

In the experiment the black and white areas were the reverse of those shown here, the grid being white on a black ground.

formations were simple, the invariants being the total width of the figure, the distances between vertical lines, and the height of the vertical lines. The variant was the angle made with the horizontal by each of the quasi-horizontal lines. The angles for the horizontal lines of the undistorted grid and the quasi-horizontal lines of the five transformations were respectively 0° (Cond. A), 2½° (Cond. B), 5° (Cond. C), 10° (Cond. D), 20° (Cond. E), and 40° (Cond. F). The induction designs were projected from slides onto a beaded screen placed in such a position that light from the design would be reflected into the right eye by a mirror placed at a 45° angle in front of the eye. The distance from the center of the projected design to the center of the mirror was 39½ in. which, added to the distance from the center of the mirror to the eye (2½ in.), gave a total distance from design to the eye of 42 in., the same as the distance from the test object to the eye. The same degree of accommodation and the same angle of convergence were present regardless of whether the test object or the induction design was being shown, except that the size of the induction design might have caused slight changes in accommodation and convergence, since all parts of the image, being in a plane, were not at the same distance from the eye.

The primary purpose of the experimental design was to provide data from which the relative after-effects, if any, of the five geometric transformations could be calculated. The design controlled (*a*) order of presentation of condition, (*b*) temporal position in the series of stimulations, and (*c*) differences between *S*s.

A 6 x 6 Latin square, replicated six times to unconfound the individual differences and order of presentation, was used. This design is shown in Table 1. The 36 *S*s were assigned to the various orders at random.

TABLE 1
PLAN OF EXPERIMENT

| *S*s | Order | Temporal position | | | | | |
|---|---|---|---|---|---|---|---|
| | | α | β | γ | δ | ε | ζ |
| 1- 6 | I | A | D | F | C | E | B |
| 7-12 | II | D | C | B | A | F | E |
| 13-18 | III | F | B | D | E | A | C |
| 19-24 | IV | C | A | E | D | B | F |
| 25-30 | V | E | F | A | B | C | D |
| 31-36 | VI | B | E | C | F | D | A |

Greek letters heading the columns represent the six different temporal positions of the experimental conditions presented to each *S*. The Roman numerals at the left of each row represent the six different orders of presentation. The six experimental conditions, occurring as six geometric transformations, are represented by Latin letters appearing in the cells of the squares. The letters were assigned to the square in random order within the limits imposed by the provision that a given letter must appear once but no more than once in each row, and once but no more than once in each column.

The *S* was brought into the room, seated at a chair in front of the apparatus, and asked to place his forehead against a concave, padded hood which approximately fitted the contour of the face and held the head in a fairly steady position. During the time *S* was being placed in position a 7-w. bulb was left on but so arranged in a reflector that only the hood, the chair, and the immediate surroundings were dimly visible. As soon as *S* was comfortably adjusted the light was turned out and *S* allowed 5 min. for partial dark adaptation. A pre-test was then given which was conducted as follows:

> You will see two different kinds of things during this experiment. One will be this (induction design was turned on manually for about 3 sec.). When you see that you don't have to do anything except look at it. The other thing you will see is a line or bar like this (test object was exposed for about 1 sec.). You got a good look at it that time, but in the experiment it will be visible for only .05 sec. like this (brief exposure). Did you see it? (Pause for affirmative answer from *S*.) The line will either be slanted *up* to the right like this (U10° exposed briefly) or *down* to the right like this (D10° exposed briefly). Sometimes it will be slanted much less than it was then—more like this (U4° exposed briefly)—that one was *up* to the right. Here is one *down* to the right (D4° exposed briefly). Because the line is visible for such a short time, a gong will sound just before the light goes on, like this (gong, .5 sec. dark interval, D4°; pause; gong, .5 sec. dark interval, U4°). If the line is slanted *up* to the right just say "up." If the line is slanted *down* to the right just say "down." Let's try a few for practice—ready? (Four exposures preceded by gong just as they appear in trial series, order DUUD). Here are some more (series UDDUUD). That's fine; now just look at this for a while (induction design for 2 min.). Ready? Be sure to guess out *loud* whenever you see the line.

In preliminary experiments the induction design had been presented for 8 min., but such a long period could not be used when *S* was required to observe six different induction designs (successively presented) and make 40 responses during each. Because of the short original induction period used here (2 min.), reinforcement periods of 3 sec. each were added. The induction design interval was followed by a 2-sec. dark interval to permit any after-lag to disappear. A warning gong sounded 1.5 sec. after the dark interval started, and .5 sec. later the test object was presented for .05 sec., tilted at the first of the 40 possible angles of tilt. The interval of .05 sec. was the shortest time under such low intensity of illumination in which pre-test *S*s could see the test object well enough to guess its tilt with confidence. Longer intervals were avoided in order to decrease the effect of the test object itself on *S*'s phenomenal horizontal. A later analysis of the responses indicated that no such *intra*-serial effects occurred to a measurable degree.

The exposure interval was followed by another dark interval of nearly 1 sec. during which *S* usually responded. The response was recorded while the induction design was again exposed for a 3-sec. reinforcement period and the cycle repeated with the test object in the second of the 40 test positions.

After the 40 exposures (i.e., one revolution of the cam controlling the test-object shaft) the motor was stopped and turned back by hand slightly so that a new series would start with a full dark interval. The *S* was then allowed to stand in the hall or on the steps outside the building for about 4 min. to rest his eyes and, more important for the purpose of the experiment, to reduce inter-serial carry-over from one experimental condition to the succeeding one.

During the rest period the slide for the next induction design was placed in position. The *S* was then brought in for the second experimental series which duplicated the first except for the induction design used. The instructions were given only before the first series. The dark-adaptation period before all series but the pre-test was only 1 min. A longer time between series would have been desirable, but it was felt that *S*s could not be asked to take part in an experiment lasting much over an hour, particularly when the experiment involved making 240 difficult judgments.

## C. RESULTS

During each trial series (40 exposures following a given induction design) *S* made 40 responses. The number of *"U"* responses was then used as a score representing the amount of after-effect of each induction design. Since nine test objects were objectively up, one horizontal, and 30 down, an *S* would be expected, *a priori,* to make 30 *"D"* responses and either 9 or 10 *"U"* responses, depending upon whether he called the objectively horizontal test object *"U"* or *"D."* Since most of the judgments were very difficult, however, we would actually expect extra-experimental variables to cause some variation around the *a priori* mean score of 9.5 *"U"* responses. There also was the possibility of constant errors favoring either *"U"* or

*"D"* responses. Since one of the trial series of the six given each *S* involved an undistorted induction design, a control group mean score based upon the 36 scores made on that trial series could be computed. The control group mean score was found to be 7.9, showing a slight but significant tendency to give fewer than the (*a priori*) expected number of *"U"* responses.

Since there were 36 *S*s, each presented with 240 test objects, a total of 8,640 responses were recorded. Actually there were 17 failures to respond for one reason or another. The 17 omitted responses were filled in on the basis of the responses made to the other test objects in the same trial series.

The 216 scores (36 *S*s x 6 conditions) were first subjected to an analysis of variance. On the basis of the (independent) estimates of the population variance *F* tests were made to determine the significance of the four controlled variables. The results of the tests of significance, in decreasing order of significance, are as follows: experimental conditions (Latin letters), significant beyond the .001 level; *S*s (Arabic numbers), significant beyond the .001 level; orders of presentation (Roman numerals), significant beyond the .01 level; temporal position (Greek letters), not significant.

Nearly all investigators in this general area have noted great differences between *S*s (8). Wapner and Witkin suggest the possibility that a tendency to be influenced by the field or to remain independent of it may be a consistent characteristic of the individual. In their experiments a distorted field was involved but no perceptual after-effects. Bales and Follansbee, who were measuring after-effects of curved lines, found them to be entirely absent in some *S*s (1). The present experiment, which is concerned with both after-effects and distortions of the field, indicates very considerable individual differences. The criticism might be made that *S*s whose scores indicate no significant differences between experimental conditions may have been merely answering at random, perhaps with their eyes closed. This interpretation is rendered unlikely by an inspection of the scores which show that the responses of practically all *S*s showing non-significant score differences under different experimental conditions were fairly consistent from trial series to trial series for a given position of the test object, but appear to be unaffected by the induction design.

The significant differences found between different orders of presentation undoubtedly indicate that perceptual after-effects bridged the inter-serial rest period of four minutes and had slight but measurable effects on the next trial series. This interpretation is strengthened by the fact that mere temporal position of a given experimental condition in the group of six trial series had no effect on the responses.

Since the difference between scores obtained under experimental conditions was the primary concern of the experiment, a more detailed analysis of these differences was undertaken. Since the over-all significance has been established, a comparison of the mean obtained under any one condition can be made with the mean under any other condition. The six conditions represent six quantitatively different amounts of change of a single variant. It would be expected, therefore, that the after-effects would be some kind of function of the amount of distortion. Previous experiments by nearly all investigators of after-effects have shown that the effect tends to rise to a maximum at some point and to decrease after that point has been passed. Gibson, for example, reported that the after-effect is greatest (1°-2°) for a line tipped at about 10°, falling to zero at no tip and at 45° (3). The six distortions used in the present experiment show progressively greater amounts of perceptual after-effect from 0° through 2½°, 5°, and 10°, the difference between successive mean scores significant in every case. At a distortion of 20° the mean score is lower, though not significantly lower, than the mean score under 10° distortion. It is, however, significantly higher than that of the control group. The curve of Figure 3, drawn through the six mean scores, shows the relative effectiveness of the six experimental conditions in producing after-effects.

## D. Discussion

The experiment shows most directly that if all figured parts of a visual field are systematically distorted (by a certain geometric transformation) an *S*'s phenomenal horizontal will be influenced in the direction of the coordinates of that system, the effect will be measurable after the visual stimulus field and visual after-effects have disappeared, and the amount of the effect will be a curvilinear function of the amount of distortion.

Several additional inferences from the data may be made, one regarding the factors mediating the after-effects. The fact that the after-effects are measurable by test objects presented to the left eye when the induction design was presented to the right eye implies that they must be mediated post-retinally. This could conceivably be done through postural or motor mediation, perhaps minimal verbal formulae, the amount of tension in certain muscle groups or the sub-vocal words in verbal formulae representing the amount of distortion of the induction design being carried over to act as a frame of reference for the test object. This interpretation is a logically possible one, though several facts make it improbable. The fact that large amounts of distortion were less effective than moderate amounts would hardly be expected. It is possible that the curve would tend to level off at higher dis-

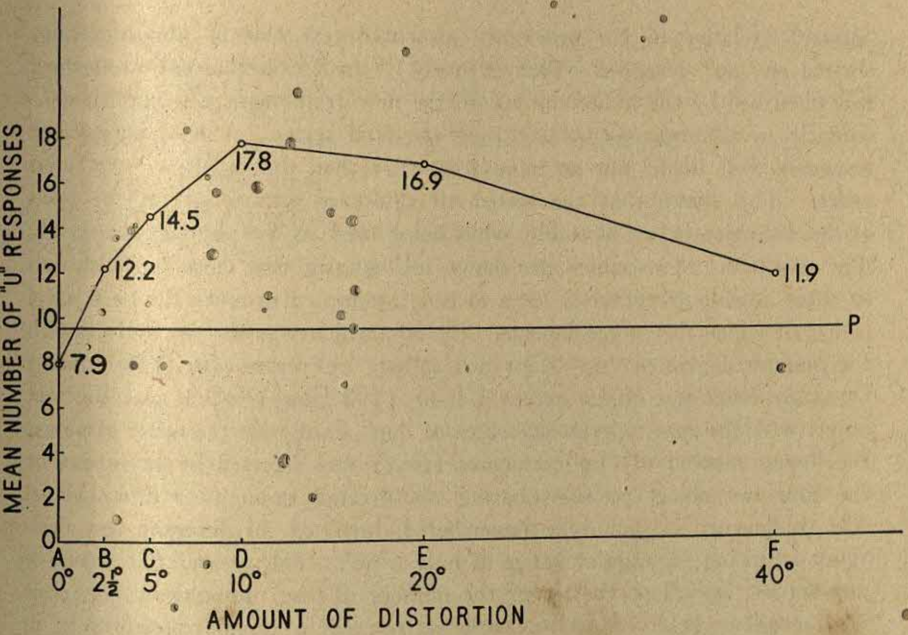


FIGURE 3
CURVE REPRESENTING AMOUNT OF PERCEPTUAL AFTER-EFFECT AS INDICATED BY RESPONSES OF SUBJECTS UNDER SUCCESSIVELY GREATER AMOUNTS OF DISTORTION (A, B, C, D, E, F)
Height of ordinate represents mean number of *"U"* responses. Line *P* represents represents a priori prediction (9.5 *"U"* responses).

tortion, but a falling curve after a certain point is reached would seem unlikely. Several (eight) *S*s indicated by their remarks that they did not realize that induction design *B* was distorted, yet in every case but one they showed a higher score on their responses to *B*. The one exception showed a generally erratic pattern of response. The strongest bit of evidence against motor mediation is found in the highly significant differences between scores obtained under different orders of presentation. Since *S* left the room during a four-minute rest period between each two successive trial series, a postural or motor set representing the previous induction design would probably disappear, whereas a non-motor after-effect, such as a visual or perceptual after-effect, is more likely to be retained.

The responses made to the varying positions of the test object during a given trial series show that nearly all *S*s were very consistent. If a test object was responded to by "up" (regardless of its objective slant), subsequently presented test objects which were objectively more "up" or less "down" (relative to the previously presented test object) almost always elicited an "up"' response. That is to say, if an *S*'s phenomenal framework was disturbed by the induction design, the new framework was applied consistently to all responses made during the trial series. A few inconsistent responses were made, but no more frequently than during the control trial series. This shows that the system of reference established on the basis of the distortion is just as stable, while being used, as the normal framework. The consistency of responses also shows, incidentally, that most *S*s (with two or three notable exceptions!) seem to have attempted to make the best judgments of which they were capable. U1°30′, which was the test object with the greatest degree of "up" slant was called "up" more often (184 times) than any other test object except U1°20′ (192 times) which was the test object with the next to greatest degree of "up" slant. At the other extreme, the lowest number of "up" responses (four) was received by three out of the four test object positions having the greatest amount of down slant. The judgments, it should be remembered, involved 40 different test positions occupying an angular range of only 6°30′. The greater the objective amount of "up" slant the larger the number of "up" responses made; this correspondence held very well over the entire range of angular positions.

## E. SUMMARY

The purpose of the present study was to investigate certain after-effects of exposure to distorted visual fields. It was found that *S*'s phenomenal horizontal was influenced in the direction of the coördinates of the distorted system, that the effect was measurable after the visual stimulus field and visual after-effects had disappeared, and that the amount of the effect was a curvilinear function of the amount of distortion.

## REFERENCES

1. BALES, J. F., & FOLLANSBEE, G. L. The after-effect of the perception of curved lines. *J. Exp. Psychol.,* 1935, **18**, 499-503.
2. GIBSON, J. J., & MOWRER, O. H. Determinants of the perceived vertical and horizontal. *Psychol. Rev.,* 1938, **45**, 300-323.
3. GIBSON, J. J. Vertical and horizontal orientation in visual perception. *Psychol. Bull.,* 1934, **31**, 739-740.
4. KOFFKA, K. Principles of Gestalt Psychology. New York: Harcourt Brace, 1935.
5. ———. Some remarks on the theory of color constancy. *Psychol. Forsch.,* 1932, **16**, 329-353.
6. NEAL, E. Visual localization of the vertical. *Amer. J. Psychol.,* 1926, **37**, 287-291.
7. THOULESS, R. H. Phenomenal regression to the real object. *Brit. J. Psychol.,* 1931, **22**, 1-30.

8. WAPNER, S., & WITKIN, H. A. The rôle of visual factors in the maintenance of body balance. *Amer. J. Psychol.*, 1950, **63**, 385-408.

9. WERTHEIMER, M. Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.*, 1912, **61**, 161-265.

10. WOOD, H. C. The "haunted swing" illusion. *Psychol. Rev.*, 1895, **2**, 277-278.

*Department of Psychology*
*University of Washington*
*Seattle 5, Washington*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 95-104.

# NORMATIVE DATA ON THE RORSCHACH*

*Jewish Vocational Service, Chicago, Illinois*

WALTER S. NEFF AND NATHAN M. GLASER

## A. THE PROBLEM

Increasing attention has been paid by Rorschach investigators, in recent years, to study of the "normal" personality. At least two reasons account for this trend. First, there has been an interest in obtaining more elaborate and more comprehensive norms, with which deviant and pathological personalities may be compared. Second, the wider application of the Rorschach Test to fields of investigation remote from its original psychiatric setting has necessitated the study of non-clinical populations. The specific area of research with which we shall deal is the use of the Rorschach in vocational guidance.

There have been a number of different attempts to use the Rorschach as an aid in vocational counseling, and in employee or student selection. We have reports of the use of the test (in combination with various measures of scholastic aptitude) in predicting success or failure of students at the college level (4, 11, 12, 13, 14); the test was used as part of elaborate efforts to improve the selection of medical students (20); and clinical psychologists (8); in the setting of World War II, the Army Air Force performed intensive research to determine if the test would assist in the selection of Air Force flying personnel (7, 9); it was used in picking cadres for the OSS (15). In addition, there have been occasional reports in the literature of the use of the Rorschach in an effort to predict suitability for various occupations (1, 3, 10, 17, 18). The results of these studies are inconclusive, and the judgment of investigators in the field is that the value of the Rorschach in relation to guidance for specific occupational choice remains indeterminate.

A common difficulty in much of this research appears to be that whereas we know a good deal about the characteristics of various categories of clinical patients, our knowledge is much less extensive about the degree of personality variation, Rorschach-wise, to be met with among so-called normal people. This paper presents data which, it is hoped, makes a contribution to the growing mass of information on Rorschach patterns encountered in non-clinic populations.

---

*Received in the Editorial Office on August 2, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

## B. The Subjects

The data reported are based on the first 100 consecutive Rorschachs administered during 1949 and 1950 by the two staff psychologists of a vocational guidance agency.[1] Of the 900 persons who were counseled at the agency during the period under consideration, 100 were deemed to have serious enough personality problems so that a Rorschach examination was requested. These 100 cases comprise our sample.

The mean age of the group was 23-3, with the median at 21-11; two-thirds of the cases fall between 17 and 24 years of age. The group was well above average in education as compared to the general public (see Table 5, below). Eighty of the cases were male, 20 were female.

## C. Results

Our chief results are summarized in Table 1, which presents means and sigmas on single Rorschach variables for the three chief diagnostic categories and for the group as a whole. It is possible for us to present the data in this form since each case was individually diagnosed as it was met in course of Agency operation, during the two years covered by this study. Examina-

TABLE 1
MEASURES OF CENTRAL TENDENCY AND VARIABILITY ON SINGLE RORSCHACH VARIABLES OF VOCATIONAL SUBJECTS, ACCORDING TO DIAGNOSTIC CLASSIFICATION, AND AS A TOTAL GROUP

| Vari-ables | Normals (*N*: 50) Mean | *SD* | Neurotics (*N*: 41) Mean | *SD* | Psychotics (*N*: 9) Mean | *SD* | Total group (*N*: 100) Mean | *SD* |
|---|---|---|---|---|---|---|---|---|
| R | 42.0 | 22.0 | 40.9 | 26.2 | 37.6 | 25.5 | 41.2 | 24.2 |
| M | 5.8 | 4.3 | 4.9 | 4.5 | 1.7 | 1.5 | 5.1 | 4.8 |
| Sum C | 4.3 | 3.7 | 3.5 | 2.8 | 4.6 | 3.4 | 4.0 | 4.7 |
| C | .38 | .61 | .27 | .51 | 1.0 | 1.2 | .40 | .88 |
| CF | 3.2 | 2.3 | 2.8 | 2.1 | 2.7 | 2.6 | 3.0 | 2.2 |
| FC | 2.1 | 2.0 | 1.4 | 1.1 | .67 | .82 | 1.3 | 1.1 |
| Y | 3.9 | 3.3 | 3.9 | 3.5 | 1.8 | 1.7 | 3.7 | 3.3 |
| V | 1.7 | 1.5 | 1.9 | 1.0 | .67 | .67 | 1.5 | 2.9 |
| P | 7.5 | 1.9 | 6.8 | 2.4 | 5.2 | 2.0 | 7.2 | 2.3 |
| S | 2.9 | 2.4 | 2.8 | 2.3 | 1.4 | 1.9 | 2.5 | 3.4 |
| F+% | 76.0 | 8.5 | 67.5 | 10.8 | 56.1 | 12.6 | 71.0 | 11.5 |
| F% | 65.5 | 11.0 | 67.6 | 18.0 | 71.2 | 11.6 | 67.1 | 13.9 |
| A% | 44.8 | 11.9 | 47.0 | 17.1 | 47.4 | 21.8 | 46.4 | 13.6 |
| W% | 20.8 | 13.4 | 18.4 | 14.9 | 26.8 | 21.8 | 20.5 | 15.3 |
| D% | 72.0 | 12.0 | 69.8 | 13.2 | 63.0 | 17.9 | 71.1 | 13.3 |
| Dd% | 9.0 | 8.5 | 13.1 | 10.9 | 7.0 | 6.8 | 9.5 | 9.8 |
| Aff. R. | .61 | .17 | .61 | .23 | .62 | .26 | .61 | .23 |
| Z | 32.8 | 18.5 | 28.7 | 16.9 | 19.8 | 9.5 | 29.3 | 17.7 |

---

[1] The Jewish Vocational Service, 231 S. Wells Street, Chicago 4, Illinois. All the Rorschachs were individually administered by the authors.

tion of Table 1 shows that the data move in the direction we would expect. Those subjects whom we have diagnosed as neurotic prove, as a group, to be slightly less productive than the normals, produce fewer movement and color responses, report fewer populars, show lower form accuracy, and poorer ability to organize the stimulus material; as may be expected also, the neurotic records are more constricted than those we felt were generally free from pathology, and there is a much higher frequency of small-detail responses. When we turn to the psychotic records, of which we had only nine cases, the group differences are striking as compared to the other two sets: general productivity, human movement responses and organization scores are all lower, form accuracy is sharply reduced, the color balance is strongly unfavorable, fewer populars are seen, and wholes are perceived more frequently at the expense of both large and small details. These results tend to reassure us, that in treating our individual records by group statistical techniques on single variables (the least meaningful way to present Rorschach results), we have not obscured the basic patterns from which diagnoses are attempted.

Since we are interested primarily in the normative implications of our data, we next compared our results to those recently presented by Beck and his co-workers in a recently published study (2). In this paper, Beck presents means and sigmas on single Rorschach variables obtained from the records of 157 employees of a Chicago mail-order firm, who served as voluntary normal controls for a study of schizophrenia. Table 2 shows the mean differences on each variable between our results and those which Beck reports, for our total group of 100 cases, for our 50 normals, and our 41 neurotics. If we inspect Column 3 of Table 2, we find a number of significant differences between our total vocational group and Beck's normals for a number of the chief variables of the Rorschach (R, M, C, CF, Y, F+%, Z). If we take this result on its face value, we should infer that the patterns yielded by our group were, on the whole, richer than Beck's, with more liberation of both intellectual and affective elements of the personality, less of a tendency to be accurate and realistic, more immaturity, and somewhat greater sensitivity, anxiety, and dysphoric mood.

Being aware, however, that it is improper to compare our *total* group with the Beck normals, since our 100 cases are known to be heterogeneous with respect to clinical diagnosis, we next compared only the 50 cases, whom we diagnosed as normal, with Beck's sample. These results are found in Column 1 of Table 2. The differences found when we compare our total group with Beck's *become accentuated* when our normals are compared with Beck's normals. To the seven variables on which significant differences are

TABLE 2
COMPARISON OF VOCATIONAL TOTAL GROUP, VOCATIONAL NORMALS AND NEUROTICS WITH DATA ON A NORMAL GROUP PRESENTED BY BECK, ET AL (2)

| Variable | (1) Normal vs. Beck Diff. | (1) Normal vs. Beck *t* | (2) Neurotic vs. Beck Diff. | (2) Neurotic vs. Beck *t* | (3) Total vs. Beck Diff. | (3) Total vs. Beck *t* |
|---|---|---|---|---|---|---|
| R | 9.3 | 3.44** | 8.2 | 1.72 | 8.5 | 3.03** |
| M | 2.3 | 4.10** | 1.4 | 1.87 | 1.6 | 2.71** |
| Sum C | 1.2 | 2.64** | 0.4 | 0.00 | 0.9 | 1.70 |
| C | —0.22 | 2.20* | — 0.33 | 3.33** | —0.20 | 2.00* |
| CF | 1.8 | 6.43** | 1.4 | 2.83** | 1.6 | 6.00** |
| FC | 0.7 | 3.22** | 0.0 | 0.00 | 0.0 | 0.00 |
| Y | 1.9 | 5.10** | 1.9 | 3.35** | 1.7 | 4.70** |
| V | —0.1 | 0.00 | 0.1 | 0.00 | —0.3 | 0.00 |
| P | 0.7 | 2.57** | 0.0 | 0.00 | 0.4 | 1.46 |
| S | 1.0 | 3.33** | 0.9 | 2.25* | 0.6 | 1.54 |
| F+% | —5.8 | 2.70** | —11.8 | 6.25** | —9.3 | 5.85** |
| F% | —4.7 | 2.93** | — 2.6 | 0.00 | —3.1 | 1.79 |
| A% | —1.7 | 1.01 | 0.5 | 0.00 | —0.1 | 0.00 |
| W% | 1.0 | 0.00 | —1.4 | 0.00 | 0.7 | 0.00 |
| D% | 0.1 | 0.00 | —2.1 | 0.00 | —0.8 | 0.00 |
| Dd% | 0.8 | 0.00 | 4.9 | 2.70** | 1.3 | 0.00 |
| Aff. R. | 0.02 | 0.00 | 0.02 | 0.00 | 0.02 | 0.00 |
| Z | 10.3 | 4.50** | 6.5 | 1.43 | 6.8 | 3.24** |

**Indicates *t* at better than the 1 per cent level of confidence; *indicates *t* at better than the 5 per cent level of confidence; when *t* is under 1.00, it is shown as zero. The negative signs in the Diff. columns show that in these cases the Beck figure is larger than ours; in all other cases, our figure is larger than Beck's.

shown in Column 3, must be added five more from Column 1 (Sum C, FC, P, S, F%). To complete this analysis, we finally compared our neurotic group with Beck's, and we find here a sharp reduction in the number of significant differences. We are forced to conclude that our neurotic group is more like Beck's normal sample than is our normal group. In fact, perusal of Table 2 shows that the bulk of the difference between our total group and the Beck sample may be attributed to the features of the sample we diagnosed as normal. Possible reasons for this finding will be presented in the discussion section below.

Mindful of the methodological weakness involved in averaging single variables of the Rorschach test, we next sought some way of handling statistically the individual test patterns taken separately so that we could estimate the degree of deviation from the "normal," case by case. After experimenting with a number of methods, including Munroe's "check-list" system (14), which we adapted to the Beck scoring method, we settled upon the use of Fisher's "maladjustment score" (6). Table 3 summarizes the use of Fisher's maladjustment score on our own total vocational group, and compares this with Fisher's reported results on his three groups of normals, conversion hysteria cases, and paranoid psychotics.

TABLE 3

FISHER MALADJUSTMENT SCORES OF THE VOCATIONAL GROUP AS COMPARED TO FISHER'S OWN EXPERIMENTAL GROUPS

| | Vocational group | (Fisher's groups)<br>Normal | Hysteric | Paranoid |
|---|---|---|---|---|
| Mean | 37.4 | 36.9 | 59.7 | 85.5 |
| *SD* | 17.9 | 11.0 | 19.1 | 32.4 |
| $SD_{(M)}$ | 1.8 | 2.5 | 4.3 | 7.4 |

Table 3 tells us that the mean score of our vocational group is quite comparable to that of Fisher's normals, but that the vocational group is more variable. If we assume that a maladjustment score of 50 divides the "normal" from the "psychopathological" (50 is roughly one *SD* above the mean of Fisher's normal controls), then we find that 21 per cent of our group obtain scores above this point. Scores above 70, which lie within the range of Fisher's psychotic subjects, are obtained by 6 per cent of our group. Summing up, we conclude that the use of Fisher's technique reveals that about one-fifth of our group of vocational applicants yielded Rorschach patterns as severely "maladjusted" as those of patients seen in a psychiatric clinic. This finding points up the rôle of a clinical psychologist in a vocational guidance agency.

As the final stage in our presentation of the characteristics of this group of cases, we examined the relationship between our own clinical diagnoses, made at the time the test was administered, and the later treatment of the data with Fisher's Maladjustment Score. Our own clinical diagnoses of the 100 cases distributes as follows: approximately one-half were evaluated as being within the normal range; about 40 per cent were diagnosed as psychoneurotic; in a few cases (9, in all) there were clearly marked psychotic patterns. Table 4 describes the relation between our initial clinical evaluations of the Rorschach patterns and the "maladjustment scores" calculated after the 100 cases were completed.

Two things are clear from Table 4. In the first place, the maladjustment score means move in the expected direction, the differences between them being reliable at better than the 1 per cent level of confidence. We therefore can say, with some degree of certainty, that our diagnoses are confirmed by the use of Fisher Maladjustment Score, or contrariwise, that Fisher's score accurately reflects prior clinical diagnoses. Secondly, we note that although

TABLE 4
RELATION BETWEEN CLINICAL DIAGNOSIS AND MEAN FISHER MALADJUSTMENT SCORES FOR 100 VOCATIONAL GUIDANCE CASES

| | Normal ($N = 50$) | Neurotic ($N = 41$) | Psychotic ($N = 9$) |
|---|---|---|---|
| Mean | 28.3 | 44.0 | 64.8 |
| $SD$ | 11.5 | 15.6 | 22.1 |
| $SD_{(M)}$ | 1.64 | 2.48 | 7.80 |
| | Normal vs. Neurotics | Neurotic vs. Psychotic | Normal vs. Psychotic |
| Mean Diff. | 15.7 | 20.8 | 36.5 |
| $SD_{(D)}$ | 2.97 | 8.18 | 7.98 |
| $t$ | 5.28 | 2.54 | 4.58 |

the mean Fisher scores for each of our own diagnostic classes move in the expected direction, each mean is considerably lower than the respective means for Fisher's three experimental groups (see Table 3, p. 10). This last point will be treated in the discussion.

## D. DISCUSSION

Comparison of our vocational group with Beck's normative sample reveals that our "neurotic" group is more like Beck's normals than is our "normal" group.

In attempting to account for this result, we wish first to consider a factor to which, we believe, insufficient attention has been paid by Rorschach investigators. We refer to the influence of social status, education, and culture on Rorschach performance. It is true that, since Rorschach himself, all workers in this field agree that "intelligence" somehow affects productivity, especially on such variables as human movement, organization score, and range of content. But the extent to which the entire record may be influenced by socio-cultural variables has not been adequately studied. Characteristically, in Beck's' normative paper, a socio-economic breakdown is presented for only two of the many variables studied (organization score, and mean card productivity). In a previous study of an army sample in World War II, one of us reported that normal subjects of low intellectual attainment typically produce skimpy, flat, and mediocre records, which look more like the records of neurotic individuals than like those of normal persons.[2] In this army study, "normal" records were produced in expected quantities chiefly by people of fairly high intellectual attainment. This point gains relevance to our study when we compare Beck's total group with ours in education (see Table 5). It will be seen that Beck's group, as a whole, lags considerably behind ours in education; the difference is even more marked when we recall that a number of our vocational cases were young people who were still continuing their education, whereas Beck's were presumably all people who had finished their education and entered the labor market. It is unfortunate that Beck does not present all of his results in terms of the economic breakdown he utilized for Z-score and mean card productivity, since in these two cases the scores for his executive group are quite similar to our group of 50 normal vocational subjects. This latter finding suggests the importance of presenting normative data according to a socio-economic breakdown of the subjects involved in the study.

---

[2] Neff and Lidz (16); see especially Tables 1 and 2 of their study, and the discussion, pp. 55-56.

TABLE 5

LEVEL OF EDUCATION ATTAINED BY VOCATIONAL GROUP, AS COMPARED TO GENERAL POPULATION (CENSUS STATISTICS OF 1940), AND TO BECK'S GROUP OF NORMAL SUBJECTS

| Amount of education | Vocational group (%) | General public (1940 Census: %) | Beck normals (%) |
|---|---|---|---|
| Elementary School | 5.00 | 32.2 | 12.3 |
| Some High School | 18.75 | 23.2 | 39.3 |
| High School Graduate | 33.75 | 31.9 | 35.1 |
| Some College | 32.50 | 9.2 | 10.4 |
| College Graduate | 10.00 | 3.3 | 2.8 |

A second factor which may combine with the first in producing our results turns around the fact that Beck does not present individual diagnostic evaluations of his subjects, nor does he present his data according to any kind of diagnostic breakdown. The question, therefore, of the degree of pathology which may exist within his group of assumedly normal subjects, remains unanswered. Perhaps if we had been able to compare our 50 "normal" subjects with the richest and least pathological of Beck's sample, the differences we found in comparing our "normals" to Beck's total group would diminish or disappear. We believe that this possibility points to the desirability of presenting normative data according to some sort of diagnostic classification, even when the subjects involved are all presumed to be normal.

The second major point which requires discussion rises out of the treatment of our vocational patterns with Fisher's index of maladjustment (see Tables 3 and 4). How can we account for the fact that the mean of our "neurotic" group is only slightly higher than the mean of Fisher's normals, and the mean of our "psychotic" group is just about at the level of Fisher's neurotics, while only three or four cases in our entire sample reach the level of the mean score of Fisher's psychotic group? Several possibilities suggest themselves. First, it must be recalled that our 100 cases are an "ambulatory" population, i.e., they are functioning in some way or other as part of the general public; they are *not* clinic or hospital patients and have come to the agency for employment, not treatment. We should thus expect that they have strengths which the hospital or clinic patient has already lost. Just as there are important qualitative differences between an ambulatory schizophrenic and an institutionalized schizophrenic, so there are important differences in the life situation of a neurotic who can no longer go on without medical assistance and a neurotic who somehow manages.

A second possibility rises from the fact that the clinic terms "neurotic" and "psychotic" undoubtedly refer to a fairly wide range of disturbance, rather than to a single disease entities. Those individual cases whom we diagnosed by the use of these terms may then be regarded as lying toward the lower, or healthier, area of their respective ranges, i.e., sick enough to be considered "not clinically normal," but not sick enough to be decisively impaired in all aspects of their life-functioning.

We should recognize, however, that it is possible to attach too much importance to the absolute value of the Fisher Maladjustment Score. It was developed upon rational and empirical grounds in order roughly to quantify the degree of maladjustment indicated by the Rorschach test as a whole for a given individual. It was undoubtedly expected to have large errors of measurement but to have the virtue of distinguishing approximately between groups for research purposes. In this sense, Table 4 verifies the use of this index in distinguishing between pathological and non-pathological *trends* in groups of subjects.[3] Our scores move in the direction anticipated by Fisher. On the other hand, the absolute order of magnitude of measures of this type is far from being an experimentally precise matter.

## E. SUMMARY

Data has been presented on 100 consecutive Rorschachs, individually administered in a vocational guidance setting to individuals referred for personality evaluation. Implications of the data for the study of Rorschach "norms" are discussed.

Comparison was made of the mean scores of our vocational group with those of Beck's "normal" group, on various Rorschach variables. Significant differences were found, in the general direction of a greater degree of liberation of the intellectual and affective elements of the personality, as well as more signs of anxiety, in our vocational group. The differences found are related to socio-cultural variables, and to the possible diagnostic heterogeneity of the Beck sample.

---

[3] A recent paper by DeVos (5) reports confirmation of the absolute level of the Fisher scores for normal, neurotic, and schizophrenic patients. The socio-economic status of these patients is not, however, disclosed.

Using a "maladjustment" score, the average for the vocational group is about the same as reported by Fisher for normal subjects, but the variability of our group is greater. About one-fifth of our group of vocational applicants yielded Rorschach patterns similar to those of neurotic and psychotic subjects seen in a psychiatric clinic.

About one-half of the vocational group was initially diagnosed by the authors as "neurotic" or "psychotic" on the basis of the Rorschachs. Comparison of our cases with a typical clinic population suggests that our cases fall toward the "healthier" end of the ranges covered by these diagnostic categories.

It is hoped that these data will contribute toward the accumulation of normative information on the Rorschach.

## REFERENCES

1. BALINSKY, B. Use of the Rorschach in the selection of supervisory personnel. *Rorschach Res. Exch.*, 1944, **8**, 184.
2. BECK, S. J., RABIN, A. I., THIESEN, W. G., MOLISH, H., & THETFORD, W. N. The normal personality as projected in the Rorschach Test. *J. of Psychol.*, 1950, **30**, 241-298.
3. BLAIR, G. M. Personality adjustment of teachers as measured by the multiple choice Rorschach test. *J. Educ. Res.*, 1946, **39**, 652-657.
4. CRONBACH, L. J. Studies of the group Rorschach in relation to success in the college of the University of Chicago. *J. Educ. Psychol.*, 1950, **41**, 65-82.
5. DEVOS, G. A quantitative approach to affective symbolism in Rorschach responses. *J. Proj. Tech.*, 1952, **16**, 133-150.
6. FISHER, S. Patterns of rigidity and some of their determinants. *Psychol. Monog.*, 1950, **64**, No. 1. Pp. 48.
7. JENKINS, J. G., *et al.* The Combat Criterion in Naval Aviation. Washington, D. C. Division of Aviation Medecine, Bureau of Medecine & Surgery, U. S. Navy, NRC Committee on Aviation Psychology, Report No. 6, 1950.
8. KELLY, E. L., & FISKE, D. W. The Prediction of Performance in Clinical Psychology. Ann Arbor: Univ. Michigan Press, 1951. Pp. xv+311.
9. LEPLEY, W. M. (*Ed.*). Psychological Research in the Theatres of War. AAF Aviation Psychology Research Report No. 17. Washington, D. C.: U. S. Govt. Printing Office, 1947.
10. MANN, I., & ARCHIBALD, D. A study of a selected group of women employed in extremely fine work. *Brit. Med. J.*, 1944, **1**, 387-389.
11. MARGULIES, H. Rorschach responses of successful and unsuccessful students. *Arch. of Psychol.*, 1942, No. 271. Pp. 61.
12. MCANDLESS, B. R. The Rorschach as a predictor of academic success. *J. Appl. Psychol.*, 1949, **33**, 43-50.

13. MONTALTO, F. D. Application of the group Rorschach technique to the problem of achievement in college. *J. Clin. Psychol.,* 1946, **2**, 254.

14. MUNROE, R. Prediction of the adjustment and academic success of college students. *Appl. Psychol. Monog.,* 1948, **7**. Pp. 104.

15. OSS ASSESSMENT STAFF. The Assessment of Men. New York: Rinehart, 1948. Pp. xv+541.

16. NEFF, W. S., & LIDZ, T. Rorschach pattern of normal subjects of graded intelligence. *J. Proj. Tech.,* 1951, **15**, 45-57.

17. PIOTROWSKI, Z. A., CANDEE, B., BALINSKY, B., HOLTZBERG, S., & VON ARNOLD, B. Rorschach signs in the selection of outstanding male mechanical workers. *J. of Psychol.* 1944, **18**, 131-150.

18. PRADOS, M. Rorschach studies on artists—painters. *Rorschach Res. Exch.,* 1944, **8**, 178.

19. ROE, A. Two Rorschach scoring techniques: The Inspection technique and the basic Rorschach. *J. Abn. & Soc. Psychol.,* 1952, **47**, 263-264.

20. SHOEMAKER, H. A., & ROHRER, J. H. Relationships between success in the study of medecine and certain psychological and personal data. *J. Ass. Amer. Med. Coll.,* 1948, **23**, 190-201.

*Jewish Vocational Service*
*231 South Wells Street*
*Chicago 4, Illinois*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 105-122.

# COGNITIVE SYSTEM-PRINCIPLES OF LEVELING AND SHARPENING: INDIVIDUAL DIFFERENCES IN ASSIMILATION EFFECTS IN VISUAL TIME-ERROR*

*The Menninger Foundation, Topeka, Kansas*

PHILIP S. HOLZMAN AND GEORGE S. KLEIN[1]

## A. THEORETICAL CONTEXT

The present study tests a hypothesis, derived from a general conception of personality organization, regarding the direction of differences among individuals in assimilation effects in time-error. The broad outlines of the theory have been described elsewhere (6, 7, 10), but a brief summary is necessary here to develop the rationale of the present study.

The central concern of this point of view is in the *organization* of personality, that is, in consistencies of regulation in a person's thought and action. Its aim is to isolate principles of organization in cognitive behavior, termed *cognitive system-principles,* that will account for and predict a person's typical modes of perceiving, remembering, thinking. The hope is that definitions of such generic regulatory principles can provide useful bases for explaining individual differences in cognitive performance. Individual differences in psychophysical response are considered the outcome of "preferred" forms of cognitive regulation—"preferred" in the sense that they are the organism's typical means of resolving adaptive requirements posed by certain types of cognitive problems (though not the only such means available to it).[2]

Such principles of regulation should be observable throughout a person's

---

*Received in the Editorial Office on August 4, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present study is one of a larger research program supported by a grant from the National Institute of Health. An abbreviated version of this paper was read at the Annual Meeting, American Psychological Association, Chicago, 1951. We are grateful to our friend Dr. Robert R. Holt for a careful reading of the manuscript.

[2] Elsewhere, we have variously referred to such directive principles of cognitive organization as "cognitive attitudes" and "Anschauungen" (7). However, they are not attitudes in the usual sense of implying direction toward or away from a particular class of objects; also, the ordinary meaning of attitude fails to convey their generality. The term *system-principle* in the sense defined by Angyal (1) and von Bertalanffy (17) seems more appropriate, and this term will be used throughout this paper.

cognitive behavior, and the empirical task is to construct adaptive situations that will bring them into relief.

The present study applies a set of previously isolated cognitive system-principles—*leveling and sharpening*—to the prediction of individual differences in assimilation effects upon time-error (*TE*). As described elsewhere (8) a leveling tendency implies a dedifferentiation of the stimulus field, which may manifest itself either (*a*) through the reduction, within limits, of figure-ground distinctions, leading to difficulty in extracting embedded stimuli from larger, more cohesive stimulus organizations; or, (*b*) by assimilating new stimuli to a dominating organization, this leading to failure to recognize gradual changes in the stimulus field. In general, maximum simplicity of the cognitive field seems to be the consequence of leveling. In the *sharpening* tendency, on the other hand, maximum complexity and differentiation are sought. Figure-ground distinctions are heightened, and this is expressed in sensitivity to embedded stimuli and in tendencies to exploit differentiation to the fullest extent.

## B. The Time-Error

When an *S* compares stimulus intensities presented successively, a constant error of judgment, the time-error, commonly appears. For instance, if two equal intensities are compared, the second appearing after an interval, *S* tends to call the second brighter than the first. *TE* is in this instance *negative*. This refers to the fact that the comparison stimulus (second of the pair) and the standard stimulus are no longer *subjectively* the same; the stimulus intensity which is subjectively equal to the standard is less than the standard. (Hence, a comparison stimulus which is *objectively* equal to the standard is judged more intense.) When the second, or comparison stimulus, is judged less bright than an objectively equal standard, *TE* is *positive*. In this instance, the point of subjective equality corresponds to an actual stimulus value that exceeds the standard.

Köhler's classical studies, and later Lauenstein's, gave the *TE* phenomenon crucial status in formulations of trace theory. In general, the Köhler-Lauenstein explanations (11, 13) hold that in the successive comparison experience, the two compared stimuli do not exist as isolated sensations. The first stimulus establishes a heightened ionic concentration in the cortical field. After its cessation, this concentration or *trace* gradually assimilates toward the trace level of the intensity which intervened between standard and comparison. In the usual psychophysical experiment a dark field fills the interval. Presumably, then, assimilation proceeds toward the level of this no-stimulus state. If instead of a dark field a stimulus more intense than the compared stimuli is interpolated, *S* usually considers the second of the two compared stimuli as less bright than the first (positive *TE*). The trace of the first stimulus is said to assimilate to the persisting level or trace of the interpolated bright field. The relationship of the second stimulus to the combined trace level of the cortical field thus constitutes a "step-down" gradient. Conversely, a less intense interpolated field assimilates the trace of the first stimulus, and the second stimulus of the compared pair is judged brighter. This is the outcome of a step-up gradient between the first and second stimuli; since it is judged against the overall trace level, the second stimulus seems more intense. Lauenstein observed, then, that where the interpolated field is less intense than the compared stimuli, *TE* is negative; where the interpolated field is more intense, it tends to be positive. In short, the direction of assimilation and hence the sign of *TE* is determined by the intensity of the prevailing interpolated field (13).

## C. PROBLEM

While Lauenstein observed individual differences in his test situations, these were unexplained. The qualities of leveling and sharpening cited earlier seem relevant to the *TE* phenomenon and suggest hypotheses to account for variations among people in the potency of the assimilation tendency. Knowing a person's tendency to level or to sharpen, is it possible to predict the amount of assimilation that will be reflected in *TE?*

The following propositions were tested: (*a*) Since leveling implies the reducing of figure-ground gradients and the minimizing of differences in the cognitive field, assimilation effects will be considerable in extreme levelers. (*b*) Since the process in sharpening is to maximize differentiation, extreme sharpeners will be better able to isolate the interpolated field from the compared stimuli. For them, assimilative tendencies should be weak.

Previous experiments have attributed special importance to *TE* for the understanding of memory. In Gestalt theory it has been a proving ground for conceptions of the memory trace. More recently, Postman and Page (14) have experimentally demonstrated a close relationship between *TE* and certain aspects of learning and memory. A demonstration of the predictive power of the leveling-sharpening distinction for *TE* responses would indicate, then, the generality of cognitive controls beyond perceptual situations alone; it would help point the way toward definition of generic system-principles within an even more inclusive system-arrangement (the "personality"). In this sense, the present experiment brings within the scope of personality theory a phenomenon not previously regarded as particularly pertinent to it.

## D. PROCEDURE

### 1. *Selection of Levelers and Sharpeners*

Forty-three *S*s, about equally divided between men and women and consisting of students in a mid-western university and a psychiatric aide school, were tested on a task which in previous studies had provided the main basis for tentatively defining leveling and sharpening attitudes (4, 8). The procedure is described in detail elsewhere (4). This situation, termed "schematizing," called upon *S* to judge in inches the size of squares.

*S*s, tested in groups of three, were so seated as to be equally distant and within a narrow range of angular displacement from the projection screen. In order to give an impression of the size limits within which the squares would fall, a one-inch square (slightly smaller than the smallest actual stimulus square) and a 16-inch square were shown several times. In the test proper, 14 squares varying from slightly more than 1″ on a side to 14″ were projected singly from a film strip on to the screen. Each square appeared for three seconds after which *S* had eight seconds to record on a record sheet his absolute judgment of the size of the square. At first only the first five smallest squares were shown, each three times and in fixed random order. Then, without warning and without interrupting the procedure, the smallest square of this series of five was removed and a square larger than any seen thus far was added to the series. Each square in this new series of five squares was again presented three times. *In this way a square gradually shifted from being largest in one series of five to smallest in another.* In all there were 10 series of squares and *S* gave a total of 150 judgments. In rendering judgments in inches, *S*s were allowed to determine their own scale with whatever refinement they wished. Scale preference was itself an interesting datum but was irrelevant to the aims of this study. Our critical datum was the accuracy with which *S*s *ranked* a square within a series, i.e., how well they distinguished stimuli as these moved from largest to middle and finally to smallest in a series. As will be shown below, the measures used did not wholly depend upon *S*s' preferred scale units.

Criteria of leveling and sharpening were the same used in earlier studies. *Levelers* were those who not only were inaccurate in detecting the position of squares within a series throughout the test, but also showed a high *percentage loss* of accuracy when squares moved from the salient end positions (i.e., largest or smallest of the series) to the middle positions, in any one series. They tended to judge all squares that occupied middle positions as similar in size. *Sharpeners* were those who effectively differentiated the squares in the middle positions of each series and were also highly accurate throughout.

Of the 43 *S*s tested 11 failed to meet the criteria for *consistent* leveling or sharpening. For instance, some *S*s were initially inaccurate and continued so throughout the tests. This took the form of a low level of general accuracy, accompanied by little differential loss of accuracy as the squares shifted positions. There were others who were quite accurate in judging end positions of stimuli but lost considerable precision when judging them as they occupied middle positions. This resulted in moderate overall accuracy, but also in a high percentage loss of accuracy for middle stimuli. The 32 *S*s who remained were divided into two groups on the basis of the above criteria, the upper half *sharpeners* and the lower half *levelers*. Since we wished to test our hypotheses about the relationship between *extreme* leveling and sharpening behavior and assimilation tendencies, we selected those *S*s who fell into the upper and lower 27 per cent of *both* distributions of accuracy scores and loss of accuracy scores in the schematizing test. Twenty-seven per cent is a conventional cut-off point for fixing the extremes of a distribution (5). Thus a total of 18 *S*s remained, nine in each group.

The data for the two groups on the two selection measures are summarized in Figures 1 and 2. Figure 1 shows the per cent of accurately ranked stimuli at each series position in the extreme levelers and sharpeners, i.e., the degree to which their judgments caught the actual rank order of sizes, irrespective of absolute accuracy of scale values. Figure 2 shows the per cent loss in accuracy by our extreme subjects in judging squares as these gradually moved from largest in the series (Position 5) through other positions, to smallest in the series (Position 1). (The largest square is always in the "5" position, and the figure shows the progressive loss in accuracy in detecting the positions of a square as it moved into "4," "3," "2," and "1" positions.)

The curve of the sharpeners shows less impaired accuracy on squares in the middle positions than that of the levelers. For instance, as a stimulus changed from Position 5 (largest in a series) to Position 4 (next largest), the loss is 37.7 per cent for the levelers as compared to 7.3 per cent for sharpeners, Both groups were far more accurate in discriminating squares at the end positions of each series.

## 2. *Time-Error Procedure*

The 18 *S*s were tested for *TE* in visual brightness discrimination under three conditions, which conformed to the main outlines of Lauenstein's procedure (13): In Condition I *(dark)*, no interpolated stimulus was interposed

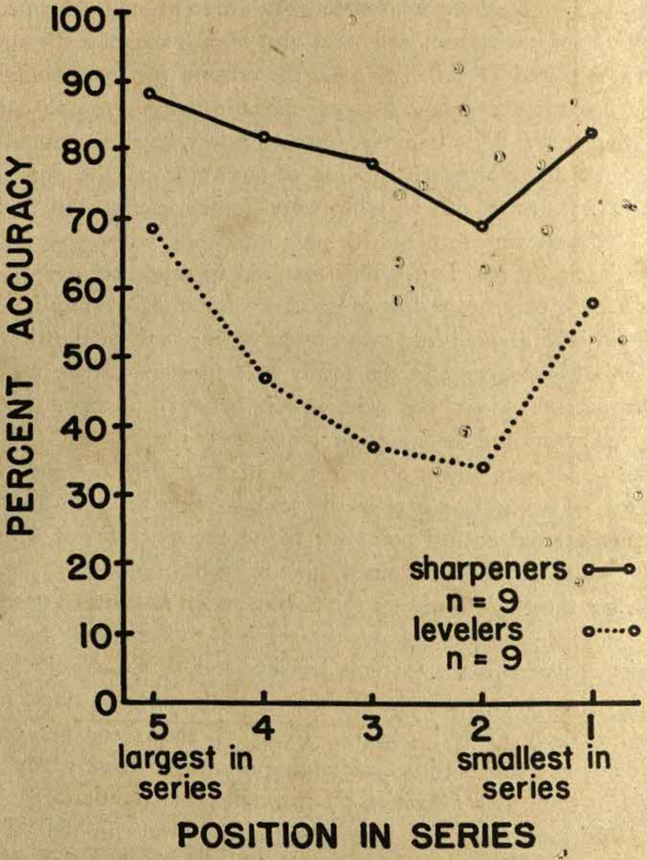


FIGURE 1
PERCENTAGE OF ACCURATE RANKINGS BY LEVELERS AND SHARPENERS WHEN SQUARES SHIFTED FROM LARGEST IN SERIES TO SMALLEST IN SERIES

between standard and comparison; in Condition II *(bright)*, a stimulus *more intense* than any of the compared stimuli was interpolated between standard and comparison; in Condition III *(dim)*, a *less intense* stimulus was interpolated between standard and comparison.

The standard stimuli were projected onto a translucent screen. Light bulbs built into the rear of the screen gave the illumination of the interpolated field. Both stimuli and interval field were projected from behind the screen. The screen was masked with black paper so that only a circular area five inches in diameter was illuminated on any one trial. Comparison, standard, and interpolated stimuli covered the same area on the screen. In each Condi-

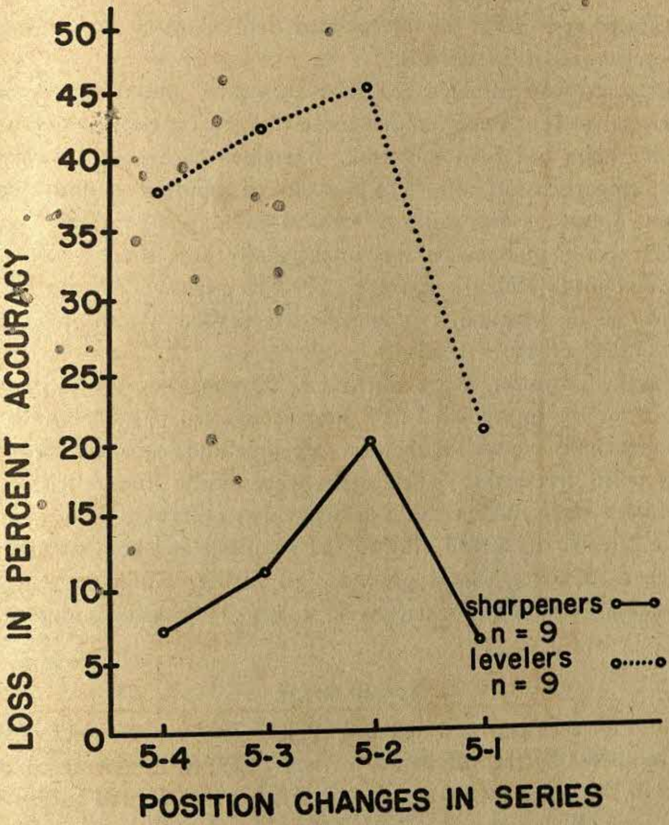


FIGURE 2
LOSS IN PERCENTAGE OF ACCURATE PLACEMENTS BY LEVELERS AND SHARPENERS WHEN SQUARES SHIFTED FROM LARGEST TO SMALLEST IN SERIES

tion, standard and comparison stimuli interrupted the interpolated stimulus which otherwise prevailed, whether dark, bright, or dim. The standard stimulus was 7 ml. in brightness and always appeared first, followed by one of five comparison stimuli. These stimuli were 3.50, 5.25, 7.00, 8.75, and 10.50 ml. Under Condition I the interpolated illumination was always the darkened screen. Under Condition II, it was 27.7 ml. Under Condition III, the interpolated illumination was 3.25 ml. The cycle of stimulation began with a 20-second exposure of the interpolated field illumination (a dark field); the standard stimulus then appeared for one second, followed by the interpolated field for 10 seconds; finally the comparison stimulus came on for one second. Twenty seconds elapsed between stimulus pairs with the interpolated stimulus taking over. Thus the interpolated field appeared for 30 seconds of the total cycle time of 32 seconds.

Stimulus pairs were so arranged that each pair preceded and followed every other pair at least once, and not more than twice; no pair was repeated until all five pairs had been projected. *S* judged each stimulus pair seven times, 35 comparisons in all. The method of constant stimuli was used throughout. *S*s used a two-category scale in judging—"brighter," or "dimmer." An experimental session was intentionally kept short (18 minutes). This took account of Köhler's warning (11) that negative *TE* can be washed out or reversed in direction if judgments are massed within single sessions ("sedimentation" effect).

The stimulus projector, a 150 watt *SVE,* 35 mm. slide projector, and the light source for the interpolated field, were connected to a mechanical timer which automatically controlled the exposure time and sequence. *S*s faced the screen six to 10 feet away. Their desks were dimly illuminated to permit them to record their judgments. The room was otherwise dark.

*S*s were directed to attend only to the standard and comparison stimuli, and to ignore the interpolated ground. To further minimize practice and "sedimentation" effects (11), one week separated the tests under each condition.[3]

## E. RESULTS

*TE* scores for each person were based upon the Point of Subjective Equality (*PSE*), computed by the summation method (19); *PSE* subtracted from the mid-point of the series gave a measure of *TE* for each *S*. All measures were expressed in millilamberts.

The data were analyzed with three points of interest in mind: (*a*) The course of *TE* under each condition. Our intention was to compare our results with trends reported by Lauenstein. (*b*) The difference between levelers and sharpeners under each condition. (*c*) The *change* in time-error from Condition II (bright interpolated field) to Condition III (dim interpolated field) in both groups. The last measure permitted a more direct comparison of the effect of the different interpolated fields on *TE* for both groups than did the separate *TE* measures for each condition. It was expected that a greater change in *TE* should occur in the leveling group from Condition II (bright) to Condition III (dim) than in the sharpening group; i.e., if levelers are more susceptible to the assimilation effect, then their *TE* should more clearly reflect the actual change of intensity of the interpolated fields than should the *TE* of the sharpener group.

---

[3] A more detailed description of the apparatus is given in (4).

The general effects of the three conditions and two groups upon *absolute level* of *TE* were analyzed through a double classification analysis of variance.

1. *Does intensity of interpolated field influence TE in the directions reported by Lauenstein?* Table 1 summarizes the analysis of variance, indicating that the estimate of variance based on Conditions is greater by far than any other estimate. Thus, the brightness of the interpolated field has a very reliable effect upon *TE*.

TABLE 1
ANALYSIS OF VARIANCE OF *TE* SCORES OF LEVELERS AND SHARPENERS UNDER THREE INTERPOLATED ILLUMINATIONS

| Source of variance | Sum of squares | df | Mean square | F | P |
|---|---|---|---|---|---|
| Groups | 1.91 | 1 | 1.91 | | |
| Subjects (within groups) | 83.87 | 16 | 5.24 | 5.29 | <.01 |
| Total between Ss | 85.78 | 17 | | | |
| Conditions | 93.06 | 2 | 46.53 | 47.00 | <.01 |
| Groups x Conditions | 10.72 | 2 | 5.36 | 5.40 | <.01 |
| Pooled Ss x Conditions | 31.78 | 32 | .99 | | |
| Total within Ss | 135.56 | 36 | | | |
| Total | 221.34 | 53 | | | |

Error term best estimate is .99

As for *direction* of the effects, shown in Figure 3, Condition I (dark interpolated field) yields an average *TE* of —.71 for all *S*s. That is, with a dark field, *PSE* was displaced .71 ml. *below* the objective mid-point of the comparison stimulus series. Under Condition II (bright field), *TE* was positive (the subjective mid-point was displaced upward 1.41 ml.), while Condition III (dim field) again yielded a negative *TE* of —1.85. In general the results confirm Lauenstein's except for the finding that the less intense interpolated stimulus (Condition III) resulted in a greater displacement of the *PSE*, and hence a *larger TE*, than did the dark interpolated field Condition, a result which closely accords with findings of Pratt and others (15, 3). The last result, however, is explicable without embarrassment to Lauenstein's interpretations and will be further considered later.

2. *Differences between levelers and sharpeners under each Condition.* The estimate of variance attributable to Groups turns out not to be significantly larger than that based on the pooled Subjects x Conditions interaction. This reflects the fact that when the results for all Conditions are combined, differences between Groups are necessarily masked by algebraic summation of negative and positive *TE*s of the bright and dim Conditions.

Evidence for the significance of leveling and sharpening is more directly

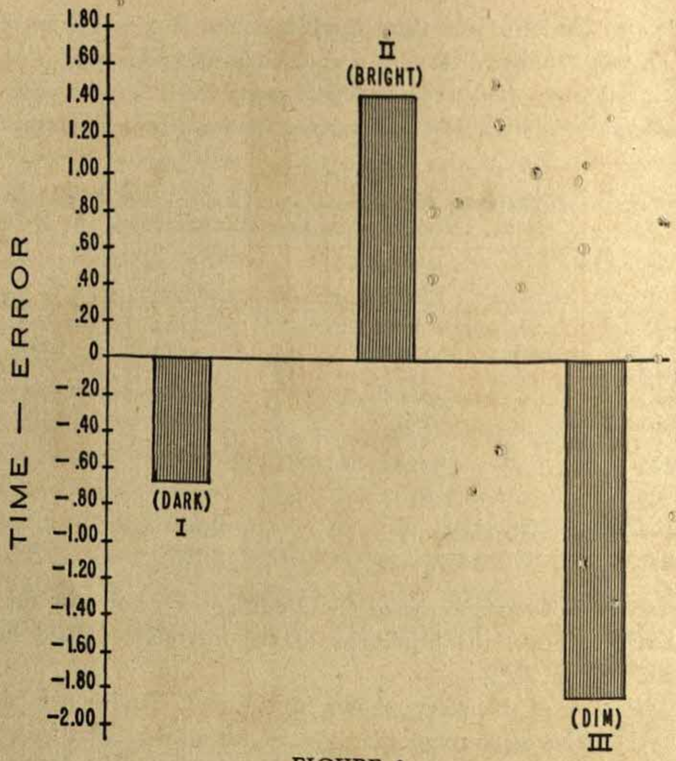


FIGURE 3
MEAN TIME-ERROR OF ALL Ss (N=18) UNDER THREE CONDITIONS OF INTERPOLATED ILLUMINATION: I (DARK), II (BRIGHT), III (DIM)

indicated by the significant *interaction* of Conditions x Groups. Figure 4 shows graphically how the Groups separated on each Condition. Differences among the Groups are all in the direction predicted by hypothesis but meet the .05 level of significance on a one-tail test only for Condition III ($t$=2.07, $p$=.03).

Variance of Subjects within Groups is also significantly larger ($p$=.01) than the pooled Subjects x Conditions variance, indicating that a considerable proportion of individual differences among Subjects is unaccounted for by the dichotomous grouping. The large variance for Subjects may reflect, in part, attitudinal variations within the leveling and sharpening Groups, some *S*s being more extreme with respect to these attitudes than others. The evidence indicates, however, that despite other as yet unknown determinants, individual differences in *TE* can be partially accounted for by the major separation of levelers and sharpeners, when duration and brightness of interpolated intervals are controlled.

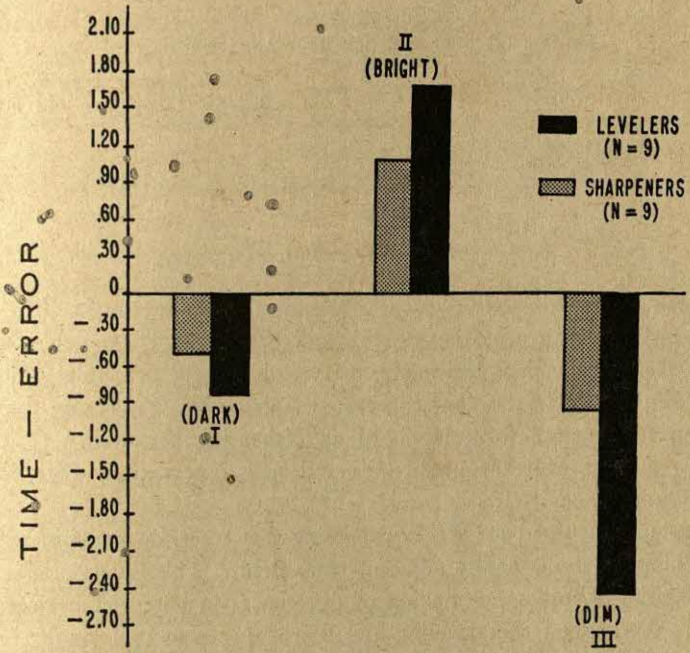


FIGURE 4
TIME-ERROR SCORES OF LEVELERS AND SHARPENERS UNDER THREE CONDITIONS OF INTERPOLATED ILLUMINATION

It is especially striking that the assimilation effect is strongest in the dim interpolation Condition, with the effect greater on levelers than on sharpeners.

One of our expectations was that a greater degree of *change* of *TE* should occur in levelers from the bright to the dim Condition. A difference score for each *S* based on an algebraic summation of his *TE* scores for Conditions II and III was computed. This measure does, of course, cumulate the separate effects of each Condition and hence maximizes the distinction between Groups, but it has illustrative value in showing the differing importance that the change in Conditions has for the two Groups. Table 2 shows that levelers have a mean difference-score that is significantly larger than the sharpeners. Clearly then, levelers and sharpeners differ in the assimilative effect of an interpolated stimulus.

## F. IMPLICATIONS

### 1. *Leveling and Sharpening in terms of brain field theory*

A theory of personality should aim to describe and interrelate, in coördinate terms, regulative principles such as leveling and sharpening, as well as

TABLE 2
DIFFERENCE SCORES OF LEVELERS AND SHARPENERS

| | Difference between Conditions II (bright) and III (dim) |
|---|---|
| Sharpeners | 2.06 |
| Levelers | 4.25 |
| Difference | 2.19 |
| t | 2.09 |
| P (one-tail) | .02 |

determinants of special effects such as assimilative tendencies in *TE*. Thus, if the decision is to accept Lauenstein's and Köhler's cortical field theory of the *TE* phenomenon, it is legitimate to try to bring the principles of leveling and sharpening within the same conceptual framework. Enlarging the special theory may accelerate formulation of an embracing personality theory and may also generate further predictions and new observations regarding the special phenomenon itself.

Lauenstein accounted for *TE* by proposing that traces do not simply "fade," but assimilate to functionally adjacent trace fields. This hypothetical process of assimilation involves a reduction of differences in potential between trace systems. We suggest that leveling and sharpening may be brought into this theory by assuming that leveling occurs in a neural field where traces are weakly bounded so that communication and fusion among traces are easy and gradients therefore reduced more readily. This state of affairs would favor assimilation to a dominant trace level. Sharpening, on the other hand, would occur in a neural system involving stronger, less easily permeable boundaries among traces, these preserving gradients more readily, thereby limiting assimilation effects. We must reserve for another report a more extended account of required assumptions concerning "trace systems," "boundaries," "fusion," and "diffusion" of traces. To summarize: leveling and sharpening would refer to different organizations of a bioelectric field—two varieties of steady states—brought about spontaneously in cortical fields differing in certain structural properties. The nature of these structural differences and their influences in affecting steady states is as yet an unclear frontier in the theory.

This interpretation has important implications, one of which is that these system-principles may act in effect as limiting factors which themselves qualify the effects of *special* directive influences in the cortical field. For instance, particular "sets" that comprise a need (e.g., those induced by a certain tissue deprivation, as in hunger) may exert differing steering effects upon the behavior of levelers and of sharpeners.

We have noted that the assimilation effect is considerably greater for the dimmer field (Condition III) than for the brighter, and exceeds even the effect of total darkness. The latter finding is puzzling, as Pratt (15) has pointed out, since logic leads us to expect that an absolutely dark field should create a steeper gradient, and hence a more marked assimilation effect. This seeming contradiction is dispelled by Lauenstein's qualifying statement regarding the limits of assimilation effects: "Only those traces that stand in concrete structural relationships with each other, can in general assimilate to each other. One can well expect that the closer the relationship the stronger the assimilation will be" (13, p. 156, *our translation*). Homologous features of the schematizing situation illustrate the point. In the schematizing situation, practically everyone was better able to distinguish the end or new, large squares. Similarly in *TE,* everyone should find it easier to disregard the intervening field when it obviously differs from standard and comparison. Thus, the interpolated stimulus should promote assimilation only up to a certain differential between interpolated stimulus and the range of compared stimuli; there is a capacity limit, and beyond it the assimilative tendency should decrease if not disappear. Now, the interpolated field brightness in Condition III (dim) was 3.25 ml., an intensity much closer to the comparison stimulus series range of 3.50-10.50 ml. than the very bright interpolated stimulus of 27.7 ml. of Condition II. Similarly, the dark interpolated field of Condition I stood far from the range of standard and comparison stimuli. By Lauenstein's reasoning, then, assimilation should also be minimal under Condition I. In short, Condition III provided the optimal conditions (in this experiment) for observing the assimilation effect.

This suggests an interesting direction of inquiry into leveling and sharpening. It would be valuable to know if differences exist between levelers and sharpeners in the *limits* within which assimilation effects can occur at all. For instance, within what range of separation of standard from comparison stimuli do assimilation effects occur in the two Groups? Also, what other structural variations in the interpolated field affect assimilation? Would alteration of *shape* of the interpolated field result in the "structural isolation" (in Lauenstein's sense) of trace systems, and reduce the assimilation effect? And would levelers and sharpeners differ in the effect of such structural changes upon *TE*? For the time being, it is reasonable to generalize as follows: Under proper conditions for assimilation (i.e., optimal difference of intensity between interpolated and compared stimuli, and optimal temporal separation of the comparison stimuli), levelers show a greater tendency than sharpeners for assimilation to develop toward the trace level of the interpolated field.

There is also the problem of cross-modal *generality* of the two cognitive system-principles. The present experiment assumes that traces issuing from stimuli in any modality have a similar status once established in the brain field. It further implies that the dynamics of leveling in visual *TE* hold also for trace processes and gradients of other modalities, say auditory *TE*, or kinesthetic *TE*. This problem has been dealt with more extensively by one of the authors (4).

### 2. *Concerning System-Organization and the Gestalt Theory of Cognition*

The nature and rôle of central influences upon perceptual organization remain unclarified in Gestalt theory. Wallach (18), a principal spokesman for the Gestalt position, assumes that an initial stage of "process formation" develops, consisting of a characteristic self-distribution of forces in the brain field, in response to a particular stimulus-configuration. In the absence of "special influences," brain-field organization then proceeds to a "natural" course of communication of this process-formation with trace. Process-formation, and communication of process with trace, have engaged the attention of Gestalt adherents, but the issue of *plasticity* of organization—how the "natural course of events" is alterable by motivational determinants such as needs, sets, and the like—remains unclarified in the theory.

Wallach maintains that motives can steer the course of cortical organization only at the phase of organization which follows process-formation itself, i.e., in the communication of process with trace. From this point of view, perception (meaning process-formation alone) is uninfluenceable by needs or special sets. Some representation, some form, or some contour must first develop for a need to act upon it. Wallach has pointed out that functionally it makes sense to assume a process-formation stage that is directly responsive only to stimulus-configuration. Such a coercive and only narrowly plastic organization serves as "a protective device" which reduces the possibility that the organism will be at the whim of every momentarily activated set.

Woodworth (20) has made the valuable observation that the Gestalt position fails to account for non-specific, situational sets which precede process-organization. Nevertheless, the heart of the matter remains a critical issue: Can variation be expected in process-formation *itself* once it does occur?

Efforts to confront Gestalt theory with the problem of personality organization are often confused by a tendency to conceive of personality determinants in narrowly motivational terms, i.e., as "sets," special "needs," "affects" or strongly valent "values," etc. Now if personality implies nothing more than such forms of selective effort linked to gratification of particular kinds, then it suffices for Gestalt theory to handle the matter of locus of "personality determinants" by relegating their influence only to the stage of process-trace communication, and there is no occasion for carping at Wallach's insistence upon a stage of coercive and uninfluenceable primary organization. A need has to influence *something; some* conversion of the stimulus configuration must precede such an influence.

Recent research suggests that a broader conception of personality organization is necessary, one that provides not only for directive influences of needs, values, and affects, but for regulatory principles that originate in the structural settings of processes which develop within the brain field. This point of view assumes that certain properties of the brain field are universals; they are "givens" in the functioning of all people. (An example would be gradient formation in the brain field which is fundamental to figure-ground differentiation.) However, variations in structural properties will result in critical modulations of all processes which involve these properties. Differences in these structural properties will affect the development of steady states and hence the nature of the configurational tendencies in the affected region. Thus, Klein and Krech (9), working from these premises, suggested certain consequences that might follow upon variations in a hypothesized property of *conductivity*. Conceptions of "ideal types," the system-principles of Angyal (1) and of Krech (12), and the present limited hypotheses of leveling and sharpening, also, in different ways, lean upon assumptions of such superordinate directive influences upon organization. Presumably, these higher-order regulative principles, reflecting differences in structural conditions of the cortical field, may also influence the effect that *lower-order* processes such as needs, sets, or affects will have upon behavior. As Holt[4] remarks, commenting on this passage, a battle fought in the air is a very different matter from a battle fought or seen on land, though the needs of the antagonists may be identical. This generally neglected point has been strongly emphasized by Rapaport (16) in his views on "ego structure."

Such structural dispositions could affect the products of *every* stage of cortical integration. In not limiting these higher-order determinants to the workings of process-trace communication, this viewpoint alerts us to expect individual differences in the manner of organization of the initial stage of process-formation itself, and to look for consequences of these differences in the later stages. This point of view would direct attention not to the presence or absence of figure-ground differentiation, but toward questions concerning the rate at which figure-ground differentiation proceeds, differences among people in this respect, qualitative considerations regarding the energy differentials involved, and the consequences that variations in all these may have for adjustive tendencies in general. In short, the "autochthonous" or self-distributive processes of the brain field, which Gestaltists have assumed to be unvarying among people, may themselves vary according to regulating principles of a different and more superordinate level than sets, needs, etc. Such over-all system qualities would constitute one aspect of the hierarchy of regulatory principles implied by the concept "personality."

---

[4] Personal communication.

It also seems profitable to accept a recent suggestion by Bruner (2) that "autochthonous *limits*" prescribe the degree of plasticity of neural field organization. From our point of view, these establish boundaries to the directive influence of system-principles upon events in the brain field. In the present experiment, the magnitude of the differential between interpolated and compared stimuli appeared to *limit* assimilation; beyond an optimal range of differences, assimilation effects tended to disappear in both types of *S*s. A capacity limit apparently was reached. Thus, the variations induced by leveling and sharpening can occur only within certain physical limits of the structures concerned.

We are, in short, suggesting a broadened setting of Gestalt theory which embraces "personality determinants" of different levels: it asserts that even at the primary stage of process-formation (as well as in subsequent stages), cognitive organization is subject to *system requirements* of the organism. In addition, particular needs or sets may affect the vicissitudes of the stimulus-process at other phases of organization. In this view, Wallach's conclusion that personality influences appear only in later stages of organization would be valid, but only for such "lower order" directive influences. An "anticipation" or a "need" may, of course, have an enduring representation cortically and must ultimately be reconceived in structural terms. The point we wish to stress now, however, is that we would expect individual differences to occur in the effects that special sets or needs will have, *as a function of the higher-order system-principles which also govern cortical events.* For instance, if we consider leveling and sharpening as examples of such higher-order controls, we are led to expect that specific needs, such as hunger or thirst, should affect the cognitive behavior of levelers and sharpeners differently.

## G. SUMMARY

The present study demonstrates that organizing principles of leveling and sharpening account in part for individual differences in assimilative effects of interpolated stimuli upon time-error (*TE*) of visual brightness.

1. In general, *TE* tended to be positive when an interpolated field of greater intensity than standard and comparison stimulus was interposed. When the interpolated field was less intense than the compared stimuli, *TE* was negative. *TE* was significantly greater for a dim interpolated field than for either complete darkness or for a bright interpolated field. This effect was perhaps due to the greater disparity that existed between the dark or bright interpolated stimuli and the stimulus series, as compared to the difference between the dim field and the same stimuli. In general, the results were consistent with Lauenstein's (13).

2. Magnitude of the assimilation effect differed for levelers and sharpeners: the difference between the *TE*s of bright and dim interpolated conditions was significantly smaller for sharpeners than for levelers.

The present experiment tends to support the hypothesis of two system-principles of organization in cognitive functioning. In demonstrating their generality upon the *TE* phenomenon it further illustrates the relevance of "neutral" psychophysical situations to personality theory.

## REFERENCES

1. ANGYAL, A. Foundations for a Science of Personality. New York: Commonwealth Fund, 1941.
2. BRUNER, J. S. One kind of perception: a reply to Professor Luchins. *Psychol. Rev.*, 1951, **58**, 306-312.
3. GUILFORD, J. P., & PARK, D. S. The effect of interpolated weights upon comparative judgments. *Amer. J. Psychol.*, 1931, **43**, 589-599.
4. HOLZMAN, P. S. Cognitive attitudes of leveling and sharpening in time-error assimilation tendencies. Unpublished doctor's dissertation, University of Kansas, 1952.
5. KELLY, T. L. The selection of upper and lower groups for the validation of test items. *J. Educ. Psychol.*, 1939, **30**, 17-24.
6. KLEIN, G. S. Adaptive properties of sensory functioning: some postulates and hypotheses. *Bull. Menninger Clin.*, 1949, **13**, 16-23.
7. ———. The personal world through perception. *In* Blake, R. R., & Ramsay, G. V. (Eds.), *Perception: An Approach to Personality.* New York: Ronald 1951. (Chap. 12.)
8. KLEIN, G. S., & HOLZMAN, P. S. The "schematizing" process: personality qualities and perceptual attitudes in sensitivity to change. *Amer. Psychol.*, 1950, **5**, 312.
9. KLEIN, G. S., & KRECH, D. Cortical conductivity in the brain-injured. *J. Pers.*, 1952, **21**, 118-148.
10. KLEIN, G. S., & SCHLESINGER, H. J. Where is the perceiver in perceptual theory? *J. Pers.*, 1949, **18**, 32-47.
11. KÖHLER, W. Zur theorie des sukzessivvergleichs und der zeitfehler. *Psychol. Forsch.*, 1923, **4**, 115-175.
12. KRECH, D. Dynamic systems, psychological fields and hypothetical constants. *Psychol. Rev.*, 1950, **57**, 283-290.
13. LAUENSTEIN, O. Ansatz zu einer physiologischen theorie des vergleichs und der zeitfehler. *Psychol. Forsch.*, 1932, **17**, 130-177.

14. Postman, L., & Page, R. Retroactive inhibition and psychophysical judgment. *Amer. J. Psychol.*, 1947, **60**, 367-377.
15. Pratt, C. C. The time-error in psychophysical judgments. *Amer. J. Psychol.*, 1933, **45**, 292-297.
16. Rapaport, D. Organization and Pathology of Thought. New York: Columbia Univ. Press, 1951.
17. Von Bertalanffy, L. Problems of Life. New York: Wiley, 1952.
18. Wallach, H. Some considerations concerning the relation between perception and cognition. *J. Pers.*, 1949, **18**, 6-13.
19. Woodworth, R. S. Experimental Psychology. New York: Holt, 1938.
20. ———. Situation-and-goal set. *Amer. J. Psychol.*, 1937, **50**, 130-140.

*Research Center for Mental Health*
*New York University*
*Washington Square*
*New York 3, New York*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 123-132.

# THE CONTINUATION OF TAPPING SEQUENCES*

*Department of Psychology, Santa Barbara College, University of California*

ROBERT M. GOTTSDANKER[1]

## A. INTRODUCTION

In two earlier studies (3, 4), the writer investigated the continuation of adjustive movements. The *S* had the task of continuing his tracking motion after the target he was following had disappeared from view. The degree of accuracy exhibited by this *prediction-motion* led to the conclusion that it could be a means for accomplishing typical tracking. When the previous pattern did not present a single rate of motion but rather had positive or negative acceleration, the acceleration was not continued. The *S* instead appeared to have continued at some kind of average of the preceding rates. This was the case even for those *S*s who were aware of the acceleration. The major characteristics of prediction-motion were found not to depend upon the specific sensory information available during continuation. The same reaction to acceleration occurred when the continuation was made with the eyes closed as with the eyes open. High consistency of individual differences was found in prediction-motion both with and without vision. Reliability estimates for six 6-sec. trials were in the neighborhood of .85 to .90.

The present study was designed to find whether the results could be generalized beyond the particular perceptual-motor arrangement previously employed. Consequently, the attempt was made to retain as much as possible of the "central" or predictive aspect of the task but to make the "peripheral" or stimulus-response aspects completely different. Instead of there being a moving target, there was a succession of sounds. Instead of continuously moving his hand horizontally, the *S* periodically tapped a key. However, the same equations were used to determine the timing between sounds as had been used for setting the rates of target motion.

---

*Received in the Editorial Office on August 10, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was performed under contract at the U. S. Navy Electronics Laboratory. The opinions expressed are those of the writer and do not necessarily represent the official opinion of the Laboratory.

## B. EXPERIMENTAL DESIGN

### 1. *Subjects*

Eighteen *S*s were used, including 14 men and 4 women. Ages ranged from 17 to 32. Twelve of the *S*s were college students and the other 6 were research and clerical workers at the laboratory.

### 2. *Apparatus*

Sequences of popping sounds were presented by a magnetic tape play-back. The *S* responded by tapping a telegraph key. Stimuli and responses were recorded as pips on a two-channel Brush recorder.

### 3. *Procedure*

The *S* was instructed to synchronize his key taps with each series of sounds and to continue tapping after the stimulus period ended. The instructions were to "try to press the key each time one *ought* to sound."

### 4. *Stimulus Patterns*

Three sequences of sounds were used. In one the "pops" occurred every half second. In the second sequence that rate was reached by positive acceleration; in the third it was reached by negative acceleration. Henceforth, the three patterns will be referred to as the *zero, plus,* and *minus* patterns, according to their accelerations. A record of one run on each pattern is reproduced in Figure 1. The pips on the top line of each chart are stimuli and those on the bottom line the responses. On the zero pattern, there were 21 stimuli in 10 sec.; on the plus pattern, 11 stimuli in 10 sec.; and on the minus pattern, 21 stimuli in 5 sec. The *S* was allowed to continue unguided for 10 sec. The patterns were analogues of the target-motion patterns described in the earlier studies.

### 5. *Trials*

There were 24 trials in which each pattern appeared eight times in scrambled order. The first two runs on each pattern were regarded as practice. Analyses are based on the last six runs.

## C. RESULTS

### 1. *Measurements*

The three runs shown in Figure 1 are all for *S* No. 8. They are typical of the records obtained. The response time to the first stimulus of the series is between .15 and .30 sec. in these examples; however, in a few instances it

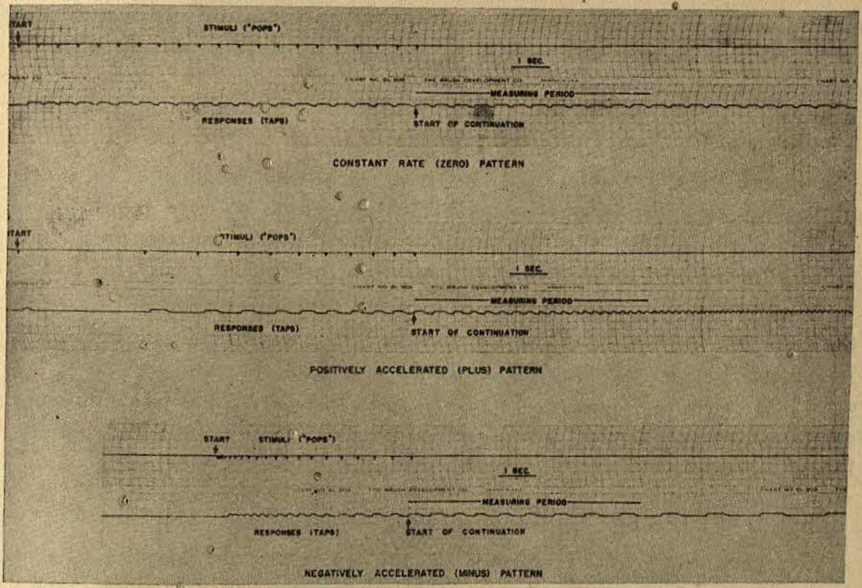


FIGURE 1

Stimulus Patterns and Typical Responses

The equations relating number of stimuli, **n**, to elapsed time, **t**, in ½-sec. units are:

$n = t$, for the zero pattern;

$n = .025\ t^2$, for the plus pattern;

$n = 2\sqrt{10t}$, for the minus pattern.

One run on each pattern is shown for *S* No. 8.

was as long as .80 sec. Synchronization of responses to stimuli generally began at the third "pop" on the zero pattern and at the fourth "pop" on the plus pattern. A "false" second tap sometimes was made about .50 sec. after the initial stimulus on the plus pattern, as the *S* apparently was anticipating the zero pattern. Because the initial stimuli came in such rapid sequence on the minus pattern, the first response usually occurred at about the time of the sixth or seventh "pop," i.e., from .20 to .31 sec. after the onset of the series. At this time the tapping speed of the *S*s equalled the ongoing stimulation rate of 6 or 7 stimuli per second. This value is in accord with previous findings on maximum tapping rate (2, p. 42). Synchronization of responses to stimuli was good by the eighth "pop" in almost every case.

On each run, tapping rate was determined over the 6-sec. interval immediately following the response synchronized with the last stimulus. The measuring period is indicated in Figure 1 for all three of the runs. Rates were expressed at taps per ½-sec. Maintaining the terminal rate on each pattern would thus provide a continuation rate value of *1*. Readings were made by ruler to the nearest .05 sec. Linear interpolation was used in computing rates when necessary.

### 2. *Classification of Subjects*

It immediately became obvious that some *S*s attempt to incorporate acceleration into their continuations on the plus and minus patterns while others do not. Each *S*, after the completion of his runs, was questioned on this matter. Of the 18 *S*s, 11 stated that they consistently tried to accelerate while 5 did not mention change of rate. The other two *S*s were inconsistent. In the results to be presented, account is taken of the two major approaches. Averages are given not only for all 18 *S*s as a group but also separately for the 11 *"accelerating"* and the 5 *"constant-rate" S*s. It should be emphasized that this classification was based entirely on the *S*s' statements and did not involve prior analysis of their tapping responses.

### 3. *Continuation Tendencies*

Mean continuation rates for the six runs are shown in Table 1. Along with them are presented means from the first study of continuous lateral hand motion (3), adjusted for differences in the constants of the equations of the two studies. The *required* values listed are those which would result from exact continuations of the equations of tapping or lateral motion.

As in the lateral-motion study there was no appreciable bias of the tapping groups from the required value of *1* on the zero pattern, the mean being 1.004 for all 18 *S*s. The means of 1.28 and .81 on the plus and minus patterns make it appear that their required values (1.31 and .81) also were closely met. However, the "accelerating" and "constant-rate" groups both departed from the requirements but in opposite directions. The "constant-rate" group's means (1.02 and 1.03) are not significant deviations from the value *1* (indicative of maintaining the terminal rate) toward the required values. On the other hand, the means of 1.44 and .72 for the "accelerating" group are highly significant and even *more extreme* departures from 1 than were required ($t = 6.98$, $P < .001$; $t = 6.83$, $P < .001$). Both of the tapping groups behaved differently than the lateral-motion group of the previous study, whose rate on the plus pattern was *lower* than its rate on the minus pattern.

A check was made to ascertain whether the departures from the value *1* by the "accelerating" group represented actual accelerations since they could have been due to deviant constant rates. A *rate increment* was found for each run on the accelerated patterns by subtracting the initial rate within the 6-sec.

TABLE 1

MEAN CONTINUATION RATE OF TAPPING IN PRESENT STUDY (IN TAPS PER 1½-SEC.) AND OF LATERAL-MOTION IN PREVIOUS STUDY (IN MM. PER ⅛-SEC.)

| | | Obtained rate | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Tapping *S*s | | | | | | 20 Lateral-motion *S*s | |
| | | All 18 *S*s | | 11 "Accel." *S*s | | 5 "Const.-rate" *S*s | | | |
| Acceleration pattern | Required rate | Mean | SD | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| Zero | **1** | **1.004** | .023 | **1.008** | .023 | **1.002** | .030 | **.99** | .11 |
| Plus | **1.31** | **1.28** | .26 | **1.44** | .20 | **1.02** | .06 | **.87** | .10 |
| Minus | **0.81** | **.81** | .18 | **.72** | .13 | **1.03** | .04 | **1.04** | .15 |

continuation period (reciprocal of the time between the first two taps) from the terminal rate. The required increment for the plus pattern is +.56 and that for the minus pattern is —.24. The obtained means for the "accelerating" group were +1.23 (greatly exceeding the requirement) and —.24 (exactly matching the requirement). In both cases the mean acceleration was highly significant, i.e., different from O ($t = 5.90$, $P < .001$; $t = 10.20$, $P < .001$).

The correlation coefficient between the continuation rates on the odd and even trials on the zero pattern was .74. Correction by the Spearman-Brown formula yielded a reliability estimate of .84 for all six trials. Because of the heterogeneity introduced by the different intentions of the "accelerating" and "constant-rate" *S*s, a comparable reliability estimate could not be made on the other patterns.

### 4. *Consistency and Accuracy*

For each *S* the ***average deviation*** of continuation rates from the mean rate of his six runs on each pattern was determined. The mean *AD*s of all 18 *S*s and of the "accelerating" and "constant-rate" sub-groups are presented in Table 2 along with comparable data from the lateral-motion study. The .015 mean value for the present study on the zero pattern is distinctly lower than the .071 value for the lateral-motion study ($t = 7.00$, $P < .001$). Because of the disparate *SD*s of the two groups, the method described by Cochran and Cox (1, pp. 91-93) was employed to test the significance of the difference.

On the plus and minus patterns it is again necessary to consider the "accelerating" and "constant-rate" *S*s separately. No clear difference is seen between the present "accelerating" *S*s and the lateral-motion *S*s. While on the plus pattern the "accelerating" group had the higher *AD* (.164 as compared to .080) on the minus pattern it had the lower *AD* (.062 as compared to .108). The differences between the "constant-rate" group and the lateral-motion group are sharp. The .037 mean *AD* on the plus pattern and .042 on the minus pattern are both significantly lower than the corresponding .080 and .108 values ($t = 3.91$, $P < .001$; $t = 4.71$, $P < .001$).

On the zero pattern the absolute errors from the *required* value of *1* were averaged for the six runs of each *S*. For all 18 *S*s a mean *average error* of .024 was obtained. Again using the method of Cochran and Cox, this value was shown to be significantly lower than the lateral-motion mean of .111 ($t = 6.21$, $P < .001$).

Figure 2 is a coordinate representation of the continuations on the three runs shown in Figure 1. Extreme accuracy on the zero pattern is seen as well as over-accelerations on the plus pattern and minus pattern.

TABLE 2

MEAN AVERAGE DEVIATION OF RATES (IN TAPS PER ½-SEC.) IN PRESENT TAPPING STUDY AND (IN MM. PER ⅛-SEC.) IN PREVIOUS LATERAL-MOTION STUDY

| Acceleration Pattern | Average Deviation | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Tapping *S*s | | | | | | | |
| | All 18 *S*s | | 11 "Accel." *S*s | | 5 "Constant-rate" *S*s | | 20 Lateral-motion *S*s | |
| | Mean | *SD* | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| Zero | **.015** | .006 | **.014** | .006 | **.013** | .006 | **.071** | .035 |
| Plus | **.116** | .081 | **.164** | .067 | **.037** | .018 | **.080** | .029 |
| Minus | **.058** | .024 | **.062** | .018 | **.042** | .021 | **.108** | .044 |

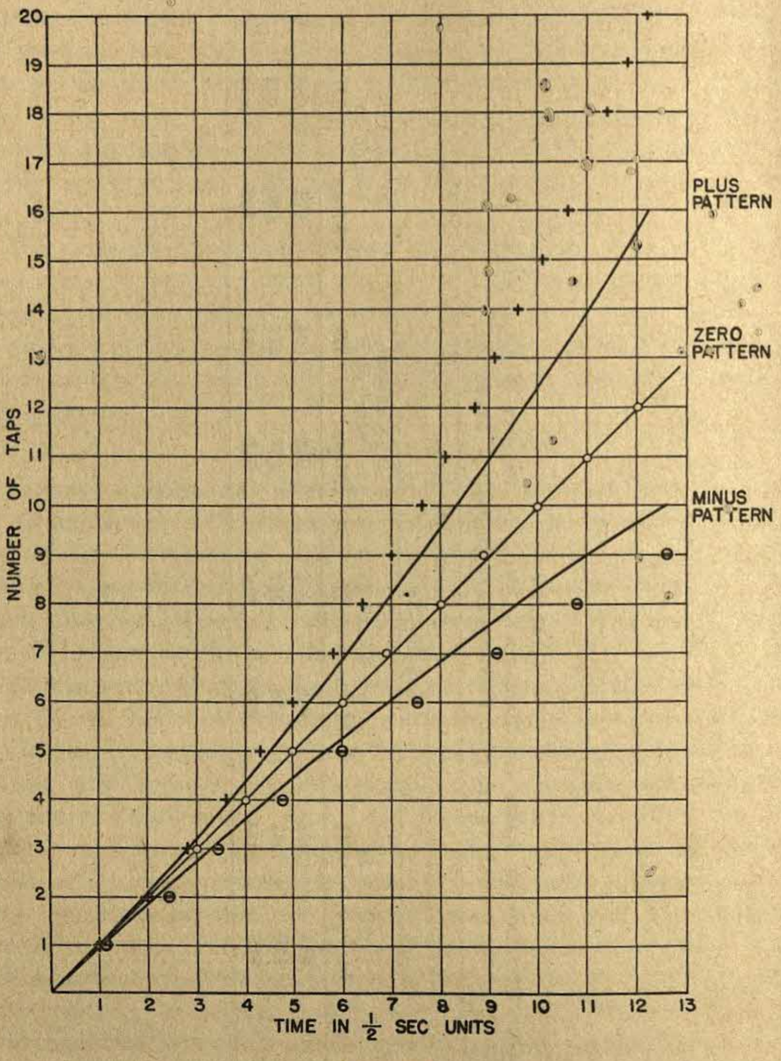


FIGURE 2

Coördinate Representation of Typical Continuations

The run on each pattern is that shown in Figure 1. Lines indicate the required continuations. The series of taps which continues each of the patterns is indicated by the series of appropriate marks, +, 0, and —. Continuation rates on these three runs are 1.62, 1.00, and .70 for the plus, zero, and minus patterns respectively. The corresponding rate increments are +1.50, 0 and —.32.

## D. DISCUSSION

In interpreting the present results in relation to those of the two earlier studies, it must be held in mind that the choice of constants in the equations of motion and tapping was necessarily arbitrary to some extent. Systematic variations of the constants would yield a more generalized answer on the characteristics of motor prediction.

No doubt whatever was cast on the previous finding that a constant rate can be continued accurately. Again, there was no appreciable bias; in addition, there was much less error per run. The mean average error in the present tapping study amounted to 2.4 per cent of the constant rate as compared with 11.1 per cent for lateral motion. Comparison of the *S*s' self-consistency from run to run in the two situations provides corroborative evidence. The mean *AD* for lateral motion was over four times that for tapping. It is not possible at this time to say why repetition rate should be more accurate than rate of continuous motion. Our understanding of even the most basic aspects of tempo is dim (6, p. 8). Nevertheless, the fact itself has certain practical implications. One past finding in particular is understandable in terms of the superiority of rapid repetitive movements over continuous undirectional motion. That is that faster rates of handwheel cranking are more accurate than slower ones (5). Tracking by very slow cranking is much like spatial positioning. Use of fast motion is tantamount to tracking by means of quick repetition of the same motion; the spatial aspect of each motion disappears and a tempo is substituted. The writer is currently engaged in research aimed at identifying the aspects of repetitive motion which give it its high level of accuracy.

The results obtained on accelerated patterns also showed differences for tapping and lateral-motion groups. Eleven out of 18 tapping *S*s incorporated acceleration into their continuations after their initial encounter with accelerated patterns. The lateral-motion *S*s did not accelerate. Further, the tapping *S*s who did not attempt to accelerate performed differently than did the lateral-motion *S*s. The tendency to smooth or average past rates which was clear-cut on the plus pattern of the lateral-motion study was not found in the present study. However, since there were only five "constant-rate" *S*s, the possibility that there may be some smoothing in the continuation of tapping sequences should not yet be dismissed.

The accelerations which were produced in the present study did not always resemble those which were required. On the plus pattern, the "accelerating" *S*s "over-did" their accelerating and also were not consistent from run to run. The stimulus-response arrangement thus provided usable information for guiding acceleration (probably because it permitted a direct comparison of time intervals) but did not necessarily lead to *successfully* accelerated performance.

Unlike maintaining a rate, accelerating is not "something that comes naturally." The *S*s who were content to maintain the terminal rate on accelerated tapping sequences showed the same superiority over the lateral-motion *S*s in regard to consistency as was found on the constant-rate pattern.

Despite the great reduction in variability of continuation rates as compared with the lateral-motion studies, the reliability estimate (.84) remained almost as high as the values previously found. The previous conclusions regarding consistency of individual difference in predictive continuation can still be regarded as valid.

## E. SUMMARY AND CONCLUSIONS

*S*s were instructed to synchronize key tapping movements with regular sequences of discrete sounds and to continue their tapping after each series ended. The continuation responses were analyzed and results were compared with those of earlier studies in which *S*s were required to maintain continuous unidirectional lateral motion. It was concluded that:

1. Repeated discrete movements may be maintained at a constant rate much more accurately and consistently than may continuous motion.

2. Acceleration is incorporated into continuation behavior more often in the tapping situation than in the lateral-motion situation, but not necessarily in an exact and consistent manner.

3. There was no evidence that the smoothing found for continuous movement occurs when *S*s attempt to maintain the final tapping rate of an accelerated pattern.

4. Individual differences in the rate at which a tapping sequence is continued are consistent as indicated by a reliability estimate of .84 for six 6-sec. trials.

## REFERENCES

1. COCHRAN, W., & COX, G. Experimental Designs. New York: Wiley, 1950.
2. GARRET, H., & SCHNECK, M. Psychological Tests, Methods, and Results. New York: Harper, 1933.
3. GOTTSDANKER, R. The accuracy of prediction method. *J. Exper. Psychol.*, 1952, **43**, 26-36.
4. ————. Prediction-motion with and without vision. *Amer. J. Psychol.*, 1952, **65**, 533-543.
5. HELSON, H. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473-479.
6. WEITZ, J., & FAIR, K. A Survey of Studies of Rhythm. Air Training Command, Human Resources Research Center. Research Bulletin 51-4, 1951.

*Department of Psychology*
*Santa Barbara College, University of California*
*Santa Barbara, California*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 133-134.

# TWO-FACTOR INTERPRETATION OF AN INSTANCE OF FEAR LEARNING*

*VA Hospital, Boston*

MERTON H. FRIEDMAN[1]

Many behavior theorists are currently exploring the relative merits of two theories of learning. One theory posits that all learning depends upon the occurrence of some form of drive reduction. This position has been most closely associated with the work of the late Clark L. Hull (2). A second theory posits that learning can best be explained in terms of a two-factor approach. Such a theory acknowledges that drive reduction can account for the acquisition of skeletal habits, but holds that drive reduction is not adequate for the explanation of other forms of learning such as those visceral-vascular reactions involved in conditioning. For the latter, two-factor theory posits that reinforcement is provided by stimulus contiguity. Such theorizing has currently been most closely associated with the work of O. H. Mowrer (3).

It is the purpose of the present paper to report an incidental finding which, it would appear, can be more adequately explained in terms of a two-factor rather than drive-reduction theory. The main investigation (1) was concerned with an experimental inquiry into the relative applicability of three theories which would attribute behavioral change following electroconvulsive shock to: (*a*) organic brain damage due to the passage of electrical current through cerebral tissue; (*b*) physiological stress due to the enfeebling effects of the violent tonic-clonic seizure induced by the shock; or (*c*) the psychological factors resulting from the fear-producing properties of the shock.

The results of the investigation here referred to indicate that there are very definite fear reactions produced in animals subjected to a series of electroconvulsive shocks. Moreover, these fear reactions are of sufficient intensity to inhibit a strongly reinforced habit originally learned in the environment in which the electroconvulsive shocks occurred. It was apparent that an originally neutral environment had assumed secondary motivating (fear-producing) properties due to its association with the electroconvulsive shocks.

It will not be necessary to enter into a detailed description of the experi-

---

*Received in the Editorial Office on August 10, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author is indebted to Dr. O. H. Mowrer for helpful suggestions provided during the preparation of this paper.

mental design other than to indicate that in that part of the experimental design where electroconvulsive shock was administered the procedure was as follows. The animals were placed in the apparatus for a period of 30 seconds and were then given an electroconvulsive shock of 0.2 seconds duration. Following the occurrence of a convulsion of full tonic-clonic proportions, the animals were immediately removed to their living cages. The shocks were administered twice a day for four days.

Thus, in such a procedure the animals were subjected to the onset of a noxious stimulus in an otherwise neutral environment. This noxious stimulus lasted for exactly 0.2 seconds, at the onset of which each animal was immediately rendered unconscious and quickly removed from the environment. Under these conditions it does not seem likely that there was an opportunity for any form of drive reduction to occur. Nevertheless, the once neutral environment, having been repeatedly paired with the onset of the noxious electroconvulsive shock, now assumed sufficient fear-producing properties to inhibit bar-pressing when the animal was once more placed in that environment under a hunger drive otherwise sufficient to motivate bar-pressing performance.

In conclusion, it would seem that the investigation here reported is an instance of fear learning, not on the basis of drive reduction but rather as a result of reinforcement provided by stimulus contiguity.

## REFERENCES

1. FRIEDMAN, M. H. Electroconvulsive shock as a traumatic (fear-producing) experience in the albino rat. *J. Abn. & Soc. Psychol.*, 1953 (in press).
2. HULL, C. L. Principles of Behavior. New York: Appleton-Century-Crofts, 1943.
3. MOWRER, O. H. Learning Theory and Personality Dynamics. New York: Ronald Press, 1950.

*1450 Commonwealth Avenue*
*West Newton, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 135-138.

# A FURTHER EXPERIMENT IN THE RECOGNITION OF EGO-INVOLVED MATERIALS*

*Department of Psychology, New York University*

M. E. TRESSELT AND MARK S. MAYZNER, JR.

## A. INTRODUCTION

In the field of law one of the principal problems concerning handwriting is the identification of samples of handwriting. Sometimes the main concern involves the recognition of handwriting where the content is unknown or sometimes the issue involves the recognition of material after a lapse of time since the original writing.

Of the several experiments on handwriting-recognition made by Inbau (3), Eagleson (2), Britt and Mensh (1), Tresselt (4) and others it becomes obvious that there are many people who cannot identify their own handwriting, even when they know the content. When the problem pertained to the recognition over a period of time, Tresselt and Levy (5) found that signatures and handwritten slogans were recognized over the period of 48 hrs., 2, 4, 6, and 8 weeks with almost equal accuracy while the identification of a square seemed to show the usual forgetting curve. Their explanation centered around the recognition as a function of ego-involvement.

The concept of ego-involvement is a relatively new one and a concept which is, at best, poorly structured. In general ego-involvement seems to indicate the "I," "me," "mine" reference for behavior activities. It reasonably follows that this personal frame of reference would delay or slow down the rate of forgetting.

However it could also be argued that the part played by the factor of overlearning in recognition should not be overlooked. Having written his own name for several thousands of times, the subject may find it just as easy to recognize his name after eight weeks as after 48 hours. The purpose of this study, then, is to determine the relative importance of overlearning and non-overlearning in relatively strong ego-involved materials and weak ego-involved materials in recognition over a period of time.

## B. PROCEDURE

There were 300 subjects in this experiment; 150 in the experimental groups and 150 in the control. The experimental groups were instructed to

---

*Received in the Editorial Office on August 21, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

write 100 times a nonsense syllable *XUHIW* which was printed on the blackboard. It was emphasized that the students were always to write in their *natural* handwriting on the paper provided. As soon as this work was done, they wrote the word once more in the center of a 3 x 5 file card. The control group only had to write the word directly on the file card without previous practice. The experimental group had to copy a nonsense drawing 100 times (the drawing consisted of a rectangle, triangle, and circle) and then put that design on a card. The control group only copied it once. Both the experimental group and the control group were further requested to write the first name of their fiance, marital partner, or steady date on another card, and on the last card they were required to draw any design they wanted consisting of a circle, rectangle, or triangle in any combination or permutation they desired.

All of the cards were collected and divided into piles consisting of (*a*) the nonsense word, (*b*) the nonsense drawing, (*c*) the given name, and (*d*) the created drawing. In order that the chance ratio be alike in all these four conditions, the four piles were sub-divided into 10 sets of 15 nonsense syllables, 10 sets of 15 nonsense drawings, 150 sets of 15 identical names, and 150 sets consisting of 15 identical created drawings.

The subjects were then divided into five experimental groups and the recognition test was given 48 hrs., or one week, or 3, 5, or 10 weeks after the initial production of the stimulus materials. The control groups were also randomly assigned to these five recognition periods. The subject was scored a failure if he did not select his own card on the first trial.

## C. DISCUSSION AND RESULTS

Figure 1 shows the percentage of correct recognition at intervals of 48 hrs., one week, 3, 5, or 10 weeks. There seems to be two distinct clusters. One cluster represents the handwriting specimens and the other cluster represents the geometrical designs. When Chi-square was obtained between these two clusters for each time-interval all points were found to be significantly different ($p$-value less than .01). These results agree in general with those previously obtained by Tresselt and Levy.

Within each cluster Chi-square tests were made for the experimental vs. the control groups at each time-interval. None of the values showed any significance. The results indicate, then, that even when a nonsense syllable is written 100 times, i.e., is overlearned, the correctness of recognition is not improved. Nor does overlearning increase the correct recognition of the geometrical designs. If the repetition of a performance such as drawing

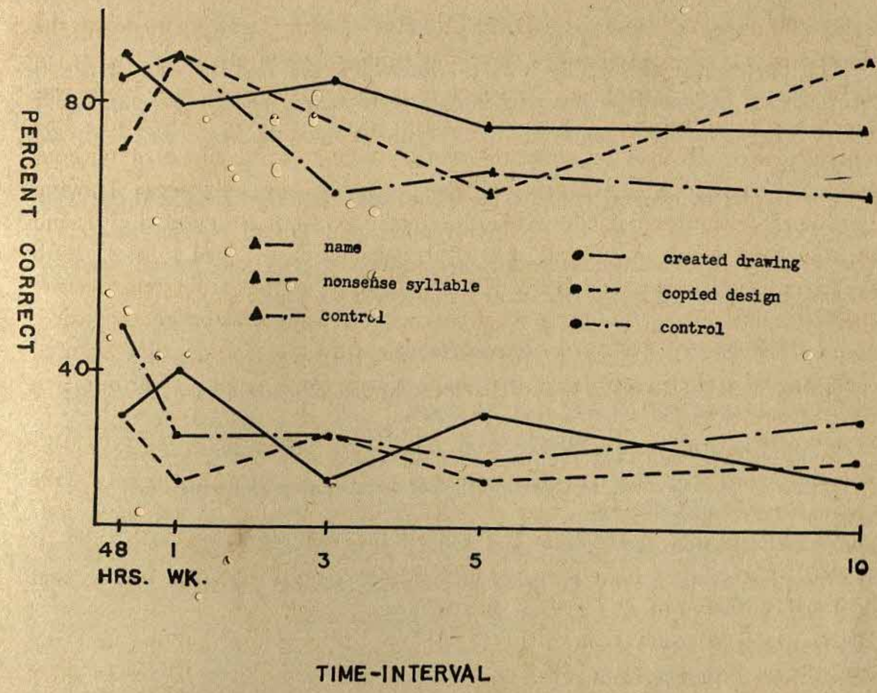


FIGURE 1
PERCENTAGE OF CORRECT RECOGNITION OF VERBAL AND GEOMETRICAL MATERIALS AS A FUNCTION OF THE TIME-INTERVAL

a figure or writing a syllable for 100 times can be called "overlearning," such overlearning plays a relatively small part in the recognition of these materials over 10 weeks.

It may be noted that recognition of the nonsense syllable is on a par with recognition of the name of the fiance, marital partner, or steady date. Apparently recognition of the nonsense syllable which one would not expect to be as ego-involved for the person as the name, functions as well possibly because of its "strangeness" and therefore salience, i.e., *XUHIW* is a seldom, if ever, written product of the subject. On the other hand, the created and copied drawings seemed to have the same weak ego-strengths.

It seems relatively reasonable to assign ego-involvement the causal rôle in explaining the results, since they do not follow the usual forgetting curve trend, but rather show a straight line relationship. Further, those materials which may be said to be relatively strong in ego-involvement are recognized significantly better than those which may be said to have weak ego-strengths.

## D. SUMMARY

The purpose of this study was to determine the relative importance of overlearning and non-overlearning in relatively strong ego-involved materials and weak ego-involved materials in recognition over 48 hrs., one, 3, 5, and 10 weeks.

The results in general support the hypothesis that the straight-line forgetting curve is a function of ego-involvement and not of overlearning; also the strongly ego-involved materials are more correctly recognized than the weakly ego-involved materials at every interval of time.

## REFERENCES

1. BRITT, S. H., & MENSH, I. N. The identification of one's own handwriting. *J. Crim. Law & Criminol.,* 1943, **34**, 50-60.
2. EAGLESON, O. W. The success of 60 subjects in attempting to recognize their handwriting. *J. Appl. Psychol.,* 1937, **21**, 546-549.
3. INBAU, F. E. Lay witness identification of handwriting (an experiment). *Ill. L. Rev.* 1939, **34**, 433-443.
4. TRESSELT, M. E. A study of the factors in the identification of handwriting. *J. Soc. Psychol.,* 1946, **24**, 101-109.
5. TRESSELT, M. E., & LEVY, B. Recognition for ego-involved materials. *J. of Psychol.,* 1949, **27**, 73-78.

*Department of Psychology*
*Washington Square College*
*Washington Square*
*New York City 3*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 139-150.

# PERSONALITY ORIENTATION: AN ORGANIZATIONAL FOCUS FOR CURRENT RESEARCH*

*Department of Psychology, University of Illinois*

RICHARD H. DANA[1]

## A. INTRODUCTION

Personality orientation is used here as a generic term to characterize the goals of broad classes of experimental endeavor. Research has been promulgated in such disparate tasks as psychophysical judgments and projective techniques. Recent symposia on *Perception and Personality* (6) and *Perception: an Approach to Personality* (3) are evidence of the crescive interest in this one-time limbo where the rigors of the experimental laboratory unite with the acumen of the clinic. Initially, this concept of personality orientation and its theoretical origins will be discussed; then, the historical antecedents and selected current research will be integrated into the proposed framework in an attempt to provide appropriate landmarks. The relevance of the concept as an organizational focus for a vast array of other research will be articulated. This approach is prototypical of the *Zeitgeist* at work in an area where many investigators implicitly acknowledge an organizing principle but few label it specifically. The ideas represented in the theoretical premises have been stated separately; their coalescence becomes one possible foundation for the proposed concept of personality orientation.

## B. THE CONCEPT OF PERSONALITY ORIENTATION

The clinical psychologist, in applying the various projective techniques, gradually becomes aware that at least one general principle guides his analyses. Rapaport (20) has suggested substituting externalization for the numerous meanings of projection. When all test results are considered as externalizations of the individual, all forms of behavior automatically become acceptable data for the clinician. However, such a formulation is useless without a standardized behavior situation against which to compare subsequent behavior samples, and without a method of classifying these

---

*Received in the Editorial Office on September 6, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The writer wishes to thank Dr. Rex M. Collier and Dr. Silvan S. Tomkins for critical readings of the paper.

diverse forms of human behavior. The concept of personality orientation can be considered as a classification device amenable to eventual concretization in terms of generalized formulas. Classification must be in terms of a frame of reference and clinical science can accept, without hypocrisy, an *S-R* approach. This means that behavior, or individual response, must be evaluated with respect to a constant, a given stimulus, a structured test, or a defined task. How much of the stimulus does the individual incorporate into his response? To what extent does he resist the prevailing environment, i.e., the particular stimulus in his response? By securing samples of the individual's performance under varied typical and controlled stimulus-conditions, generalizations can be made to performance with regard to other environmental stimulus-conditions. It is then possible to express *personality orientation* as the *relative resistance of the individual to the environmental situation.* This orientation can be measured by the extent to which protocol productions to given stimulus situations represent reproductions and elaborations (additions and distortions) of the stimulus-object or stimulus situation. A resistance ratio, or score, can then be calculated which will provide information as to the characteristic reaction of the individual to the environment. It thus becomes possible to describe individuals by the degree to which they resist the stimulus-object or stimulus situation in terms of a continuum. This continuum would constitute an operational statement: progressive increments of stimulus incorporated into response. Another manner of expressing this concept is in respect to the relative differentiation an individual has attained from the environment. Those individuals who are relatively undifferentiated from the environment resist the environmental stimuli only minimally. Therefore, they incorporate into their behavior maximal increments of the stimulus-object and are stimulus-bound. These individuals can be characterized as acceptant of their culture, conservative, centrifugal (9), and to a degree automaton conformists (12). Individuals who are highly differentiated from the environment incorporate few properties of the stimulus-object into their behavior, or response. Personal meaning is thus imposed upon the response; the result is a response-product relatively free from the stimulus-properties. These individuals tend to be creative, unaccepting of cultural prerogatives, centripetal (9), personal in orientation as opposed to environmental.

These polar individuals may be either adjusted or maladjusted depending on the extent to which they represent isolated extremes on the continuum of personality orientation. It is not necessary to spell out in detail personality characteristics of polar groups as these have appeared as the results of studies to be reported from the current literature. The exact form of the resultant curve, normal or bimodal, can only be plotted when research fills in the gaps between polar groups and attends to all individuals instead of selection of extreme cases.

## C. THEORETICAL PREMISES

The concept of personality orientation, as an integrative principle, is not proposed *in vacuo*. Certain theoretical premises are basic to this position. All human behavior is directed toward the goal of survival in an environment composed of impersonal objects and other individuals in their organized structures. Survival in our complex culture can be considered as the achievement and maintenance of biological, social, and ego constancies (homeostasis) which constitute media by which the relative differentiation from the environment is attained.

Human life, as we experience it, is actuated in a cultural milieu, and sustained by communication, by language which is a learned consequence of social interaction. To be effectively learned, the basic ingredients of communication must be simple, concise, graphic. This may explain why the logic of paired-opposites, sensed and verbalized by Aristotle, is embedded in the linguistic core of the culture and ubiquitous in thought processes and behavior.

Aristotelian logic has not only influenced scientific research but has determined the structural emphasis of Indo-European languages. As a consequence much of Western civilization has been predicated on the concept of polarities. This has occurred to the extent that all of our children are nurtured on dichotomies: love-hate, good-bad, strong-weak, big-little, masculine-feminine, responsible-irresponsible, independent-dependent, active-passive, mature-immature, etc. These dichotomous categories, the evaluative manner of thinking, represent a product of learning. These polarities are transmitted as cultural prerogatives meted by parents and their surrogates, the peer group, and larger social institutions. To an important extent this learning from which personality emerges consists in indoctrination in the logic of paired-opposites. Not only is thinking charged with this idiom, but perceptions are organized and behavior accentuated on these same premises. Polarities are in eternal conflict determined by their very incorporation into the structure of the language and life of the culture. It is no wonder then that conflict between polarities exists in every personality and is manifest in thought processes and behavior. The individual can deny the conflicts (repression) or attempt to accept, live with and resolve them in his own life. If he denies himself awareness of these polarities, he strives to maintain a precarious equilibrium in which all life is simplified to fit the natural order of the accepted dichotomy, a world of absolutes. The alternative, or relatively alternative, solution is acceptance, and implies disequilibrium, a haphazard and often chaotic attempt to reconcile the extremes; perception and action based on a relative world which is infinitely and indeterminably complicated.

All polarities are reducible to a universal dynamic: to achieve autonomy. Autonomy can be viewed in two dynamic forms: (*a*) to grow up (to attain a proper sex rôle, be adult, be the positive aspect of the dichotomy); (*b*) to achieve a unique self (resolve the dichotomy and emerge a person). Adjustment or maladjustment emerge in the contexts of both dynamic attempts to achieve autonomy. The individual who learns to perceive an ordered world and accept equilibrium is directed toward growing up, complying with the imperatives of his society. By the same reasoning, he who accepts disequilibrium of the world is directed toward the achievement of a unique self as a bastion against what Baudelaire has termed the vast *néant* of the universe. This dynamic exposition may be directed toward the environment or the person, thereby actuating the maintenance of social or ego constancies. Although the logic of polarities is incorporated by all individuals in the culture, relative influence on behavior, i.e., resolution of conflict, is related to the degree and mode of differentiation that the individual has achieved from the environment. Stagner (22) has discussed "perceptual constancy," the constancy of the self-image, and considers that the mature ego functions to render the personality more independent of the environment. Relative differentiation concerns development of constancies resultant from incorporation of a given range and quantity of environmental cues. The emphasis on social or ego constancies implies investment in ideal or real cultural values (mores). The approach to the environment is accordingly concrete or abstract, and thinking predicated in either-or as distinguished from either-and-or terms. The ultimate relationship with the environment is thus determined by the number of stimulus-cues incorporated into behavior, few or many constituting the extremes.

It is expedient, at this juncture, to summarize attributes of extreme groups on the proposed continuum of resistance to the environment. Extreme groups only are considered, areas of overlap are recognized, and the list of attributes is definitely extendible.

Those who demonstrate low resistance to the environment entertain the personality dynamic, to grow up; their learned choice is denial of polarities by acceptance of the positive pole; the result of this choice is equilibrium;

their focus is on the environment; social constancies are primary; values are ideal; world view is absolute; mode of thought is either-or; and relatively many environmental cues and relatively few personal cues are utilized in their behavior.

Those who demonstrate high resistance to the environment sustain the personality dynamic, to achieve a unique self; their learned choice is attempted resolution of polarities; the result of this choice is disequilibrium; their focus is on the self; ego constancies are crucial; values are real; world view is relative; mode of thought is either-and-or; and relatively few environmental cues and relatively many personal cues are employed in their behavior.

It is important, at this elementary stage of theorizing, to emphasize that these attributes tend to go together consistently in an individual who displays a specified degree of resistance to the environment.

Implications of this theoretical approach are obvious. Personality dynamics, when discovered by clinicians, are invariably stated in terms of the effects of a conflict between polarities. These dynamics are then incorporated into a psychoanalytic, Adlerian, Jungian or non-directive frame of reference and the personality is described. Much effort is expended searching for sets of "life-styles," or "personality dynamics," which are then considered the key to an individual. All of these formulations essentially express the same idea: the attempt to phrase man's personality in terms of relative resistance to the environment. The question asked is: how does the person react to the environment? The answer is: with varying degrees of emancipation from the environment, distance between person and objects, ability to delay immediate response. This can be thought of in terms of a continuum: relative ability to resist or not to resist the environment. It is only when conflict between polarities becomes acute, or chronic, that this relative inability to resist the environment breaks down, and the individual succumbs to environmental pressures or retreats from any further possibility of interaction. There are two alternative behavioral extremes when this occurs: incorporation of all available environmental objects (or cues), or denial of all environmental objects (or cues). Too many or too few cues for behavior remain and breakdown, dysfunction, occurs (5, 8, 23).

The obsolescence of traditional nosological categories has long been apparent. Diagnostic information, to be meaningful, must express prognosis and current labels merely stigmatize or elaborate the obvious. Comprehensive personality formulations, of which the personality orientation approach is an example, are needed to provide a basis in terms of which diagnostic information can be evaluated.

## D. Historical Antecedents

Orientation, as a personality mode, was used by Jung (14) in describing attitudes of introversion and extroversion. It was thus connected with the early dichotomization of person and environment, and the ensuing invidious categorization of "outer" and "inner" homunculi which plagued psychological thinking until recently. The person-environment dichotomy has been thoroughly dissected from various critical positions (17, 18, 19). Notwithstanding the present recognition of the limitations of the dichotomy, the attempt to separate individuals with one or the other personality mode has been omnipresent, as the review accorded this preoccupation in Vernon (24, pp. 190-213) illustrates.

The approaches of Bruner (5) and Cameron (8) stress separation of groups in terms of continua, with appropriate designation as hypervigilants and autists or schizophrenics and hysterics. "Normalcy" is inferred to be represented by some point between the extremes. "Competence and adequacy in behavior organization depend as much on what is left out as what is included" (8, pp. 58). Implicit in these approaches is incorporation of environmental cues into the behavior repertoire of the individual, i.e., relative resistance to the environment.

Tolman (23) distinguishes between broad and narrow cognitive maps. The narrow cognitive maps, product of too much or too little motivation or frustration, result in regression, fixation, displacement of aggressions, or pathology indicative of failure of cultural learning. When translation is made to the concept of personality orientation, the development of narrow cognitive maps is incorporation of too few or too many environmental cues. The individual resists the environment too much or not enough; the pathological extremes on the proposed continuum.

The most recent reaffirmation of this interest appears in Gibson (13), who distinguishes between visual world, a perceptual phenomenon, and visual field, an attitudinal product; and Boring (4), who considers System *O*, invariance, and System *R*, reduced vision, as the alternative views of the world. These represent sophisticated attempts to clarify the early crude and untenable distinctions between person and environment.

Distinctions between person and environment must be in terms of behavior in order to be meaningful. The measurement and differentiation of an attitudinal hierarchy of the "inner" man is not feasible and is not science. Scientific scrutiny must focus on measurable behavior components observed and recorded under controlled conditions. Statements about man must be operational, and treat of behavioral responses from which stimulus conditions can be predicted. Empirical research derived from this approach can not only test specific hypotheses but employ a relative resistance ratio as a standard for comparing behaviors and individuals within a consistent frame of reference. By the establishment of such a criterion clinical science can progress to the point of actually delineating the precise degree of incorporation and rejection of environmental conditions operative in personality orientation.

## E. CURRENT RESEARCH

The studies selected for elaboration have in common their dependence on the concept of personality orientation for increased interrelatedness. They employed different experimental tasks, different goals, and represent what are apparently widely separated areas of research interest. These selected researches will be discussed in order of relative complexity of experimental tasks, i.e., of stimulus materials. All of these studies place individuals at different points along the continuum of resistance to the environment. For this reason the groups isolated are approximately and not precisely equivalent. Their degree of equivalence could be ascertained by appropriate research designs as will be indicated.

Calvin and Ward (7), in an experimental *TAT* validation attempt, used certainty of reaction to a chance light-buzzer combination to divide subjects into two groups. *TAT* stories were then elicited. The high certainty reaction group was high on *TAT* needs illustrative of high resistance to the environment; the low certainty groups' *TAT* results are confirmative of diminished resistance to the environment, acceptance of the chance presentation of stimuli. This study was chosen because use of the *TAT* provided some information with regard to personality. Most studies employing simple tasks, i.e., psychophysical judgments, tachistoscopic recognition, etc., are not concerned with personality. However, the utility of the concept of personality orientation is in the to-be-tested assumption that the results of *any* task provide the relevant information from which inferences about personality can be made.

Klein (15), searching for theory germane to eventual laws of perceivers, has discerned personal styles of using perception for reality appraisal and terms these "perceptual attitudes," or *Anschauungen*. The relation of *Anschauungen* to the system of cognitive controls, or delay mechanisms, functional in personality is emphasized. Using such tasks as judgments of size of squares, phi phenomenon, Rorschach, Gottschaldt figures, etc., he finds several dimensions labelled as "leveling" vs. "sharpening," "tolerance" vs. "resistance to the unstable," and "physiognomic" or dynamic-emotive-emphatic vs. "literal." More recently, Klein (16) has discussed results of related experiments, using a thirst motivation, in which groups were distinguished in terms of control on a task demanding flexible performance. Those subjects who experienced high interference on the task, stimulus-bound, were labelled "overcontrolled" and tended to be constricted, compulsive, unadaptable, and concerned with security. Those with more "permeable control" produced more distance responses from the thirst motivation. When the results of these experiments are viewed within the conceptual focus of personality orientation, the essential relatedness of all the isolated groups becomes apparent. Thus there do not seem to be many dimensions but only *one;* that dimension which expresses the basic mode of relating to the environment, the relative resistance to the environment that the individual enunciates by his behavior. Those various *Anschauungen* groups, separated by Klein, correspond to limited ranges along the personality orientation continuum.

Witkin (26, 27), concerned principally with individual differences in perception, has completed a series of experiments which differentiate individuals emphasizing visual impression from those who emphasize bodily experience, or postural cues. A definite continuum is involved. This distinction is in terms of automatic, spontaneous experience with respect to the perception of orientation of objects, including the self, in relation to the upright in space. Thus, there appear to be individual differences in the perception of even a well-structured situation. Witkin (28), in more recent formulations, considers that dependence on visual or postural experience is ancillary to a more crucial performance factor: "It is rather a matter of ability to resist the prevailing field in separating an item (the body or any other item) from that field vs. yielding to the field or field dependence." The similarity of this expression to personality orientation is apparent and serves to accentuate the growing concern with the need for organizational concepts within the area of personality research.

At least one investigator, Frenkel-Brunswik, has recognized the relatedness of data from simple perceptual tasks and projective techniques on the same individuals. The studies under her aegis have demonstrated that tasks, or behavior, of manifold levels of complexity provide consistent personality formulations. Frenkel-Brunswik (1, 10, 11) has discussed the results of the University of California studies of ethnic prejudice. The studies of Jaensch, Kretschmer, Rorschach, Thurstone, Werner, Goldstein, and others who categorized individuals, created typologies, are characterized by Frenkel-Brunswik (11) as representative of the "perception-centered" view. Her own studies which relate personality variables (ethnic prejudice) to perception variables (tolerance of ambiguity) are considered "personality-centered." This tends to separate research areas on which recent integrative efforts have been concentrated.

A concept such as personality orientation serves to indicate the identity of "perception-centered" and "personality-centered" approaches. Resistance to the environment by the individual is not only operative but crucial to both. Ethnic prejudice is congruent with minimal resistance to the environment in which stereotypes are accepted, thinking is eschewed for conformity and security of absolutes. Relative freedom from ethnic prejudice is one expression of maximal resistance to the environment; the ability to incorporate only selected environmental cues into behavior. The generalized synonyms for high and low prejudice groups are "overconformists," "intolerants of ambiguity" or "underconformists" and "tolerants of ambiguity," respectively, depending on variables emphasized. The "overconformists" offer only minimal resistance to the environment; they are concerned with the personality dynamic of growing up; the learned choice is denial of equilibrium; the focus is on the environment; social constancies are emphasized; approach is concrete; the world appears absolute, inevitable, fixed. The "underconformists" similarly have relatively opposed and consistent relations with the environment predicated on their enhanced resistance to the environment.

In Frenkel-Brunswik's own statement the relevance of the concept of personality orientation is manifest, although the very words used suggest a dichotomy that seems to exist only in terms of structural inadequacies of the language: "This struggle between these two orientations (over and underconformity) is basic to our civilization; its individual members display these two patterns in varying proportions and changing configurations" (11, pp. 141).

Employing an art preference test with 40 male pre-doctorate students, rated high or low in personal effectiveness, Barron (2) found two types of perceptual preferences operative with equal frequency in both groups. These preferences were considered the results of "perceptual decision" to attend to aspects of the world which were either stable, predictable, ordered, or unstable, unpredictable, chaotic. The isolation of these groups, when evaluated in terms of personality orientation, is rendered more comprehensible. Barron sees behavior as a decision to accept either the positive or negative aspects of the world. The decision for acceptance of authority, of the positive aspect of the world, implies incorporation of a large number of environmental cues into behavior; individual resistance to the environment is thus minimal. To decide in favor of complexity demands a personality orientation in which resistance is heightened, disequilibrium is accepted, and the relativity of the perceived world becomes the foundation on which a unique self is predicated. In a sense the individual does have a "choice," a "perceptual decision," appearing as a direct consequence of early learning and modifiable, to an extent, by later learning, or psychotherapy. Barron's extremes, or those who made a particular decision "at its worst," are clearly schizophrenics, with the paranoid processes of stereotypy, rigidity, fear of external threat, etc., and the gross disorganization, surrender to nihilism, the epitomes.

Riesman (21), concerned with analysis of interview and projective material, posits three mechanisms of conformity: (*a*) tradition-direction; (*b*) other-direction; (*c*) inner-direction. These directions can be viewed as distinct personality orientations accepted and learned by large, overlapping segments of the population. They represent progressive accretions of resistance to the environment. Tradition-direction individuals, rare in our culture, and other-direction individuals, characterized as "radar" people, would appear to offer minimal resistance to the environment, and to incorporate the prerogatives, goals, aspirations of their culture, albeit through somewhat different media. Inner-direction individuals, the "gyroscopes," conforming to internalized norms are less dependent upon the immediate environmental cues, display greater resistance to the environment, and concentrate on ego constancy and the achievement of a unique self.

Watkins and Stauffacher (25) have designed an index of pathological thinking on the Rorschach which seems to provide some measure of personality orientation, or at least the extremes on the continuum of personality orientation. Individuals attaining a high deviation index, i.e., those who display a pathological loss or increase of distance from the Rorschach card, resist the environment too little or too much. In this manner a new method of obtaining diagnostic information by means of projective techniques is envisioned. An analogous score is being developed for the *TAT* by the present writer.

Of course, it is possible to argue that the concept of personality orientation is so vague and inclusive as to be meaningless. However, the concept, as proposed, is amenable to experimental test. High correlations should result when extreme groups as identified by *any* given task, are tested on the range and variety of tasks used by previous investigators. Feasible, too, is the establishment of a criterion score, or ratio, representing the degree of resistance to the environment. Ideally, this could provide access, in one index, to a constellation of personality characteristics. The utility of such information, obtained by one brief task or test, for psychodiagnosis and psychotherapy, is inestimable. Although it is arrant nonsense to suppose that such a shorthand procedure would be panaceatic, it is realistic to consider that awareness of an individual's personality orientation might well focalize directions in which further personality investigation should proceed.

Ultimately, however, any relevance personality orientation may have, as a concept organizing the results of diverse research, rests on the quality of theory and empirical research which emerges to sustain it. The ideas presented here are the author's view of one possible area in which comprehensive personality theory may be sought.

## F. SUMMARY

1. The concept of personality orientation, *relative resistance to the environment,* has been proposed as an organizational focus for a large body of hitherto unrelated research, and historical antecedents have been suggested.
2. Fragments of personality theory have been offered which emphasize the determinism of Aristotelian thinking on Indo-European linguistics and consequently upon behavior in our culture.
3. Several current researches, employing various tasks and goals, have been selected for exposition of the concept of personality orientation to the conclusion that *any* task or test can divulge the relevant information.
4. It is necessary to test the concept empirically and to establish a resistance score, or ratio, in order to provide indication of a syndrome of personality characteristics commensurate with a given degree of resistance to the environment.

## REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. The Authoritarian Personality. New York: Harper, 1950.
2. BARRON, F. Personality style and perceptual choice. *J. Pers.*, 1952, **20**, 385-401.
3. BLAKE, R. R., & RAMSEY, G. V. (*Eds.*). Perception: An Approach to Personality. New York: Ronald Press, 1951.
4. BORING, E. G. Visual perception as invariance. *Psychol. Rev.*, 1952, **59**, 141-148.
5. BRUNER, J. S. Personality dynamics in the process of perceiving. In R. R. BLAKE & G. V. RAMSEY (*Eds.*), Perception: An Approach to Personality. New York: Ronald Press, 1951.
6. BRUNER, J. S., & KRECH, D. Personality and Perception: A Symposium. Durham: Duke Univ. Press, 1949.
7. CALVIN, J., & WARD, L. An attempted experimental validation of the TAT. *J. Clin. Psychol.*, 1950, **6**, 377-381.
8. CAMERON, N. The Psychology of Behavior Disorders: A Biosocial interpretation. Boston: Houghton Mifflin, 1947.

9. ERIKSON, E. H. Childhood and Society. New York: Norton, 1950.
10. FRENKEL-BRUNSWIK, E. Personality theory and perception. In R. R. BLAKE & G. V. RAMSEY (*Eds.*), Perception: An Approach to Personality. New York: Ronald Press, 1951.
11. ————. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, 1949, **18**, 108-143.
12. FROMM, E. Escape from Freedom. New York: Rinehart, 1941.
13. GIBSON, J. J. The visual field and the visual world: a reply to Professor Boring. *Psychol. Rev.*, 1952, **59**, 149-151.
14. JUNG, C. G. Psychological Types. London: Kegan Paul, Trench, Trubner, 1923.
15. KLEIN, G. S. The personal world through perception. In R. R. BLAKE & G. V. RAMSEY (*Eds.*), Perception: An Approach to Personality. New York: Ronald Press, 1951.
16. ————. Measuring the regulation of motives. Paper read at the Amer. Psychol. Assoc., Washington, September, 1952.
17. KLEIN, G. S., & KRECH, D. The problem of personality and its theory. *J. Pers.*, 1951, **20**, 2-23.
18. LUCHINS, A. An evaluation of some current criticisms of Gestalt psychological work on perception. *Psychol. Rev.*, 1951, **58**, 69-95.
19. MURPHY, G., & HOCHBERG, J. Perceptual development: some tentative hypotheses. *Psychol. Rev.*, 1951, **58**, 332-349.
20. RAPAPORT, D. Projective techniques and the theory of thinking. *J. Proj. Tech.*, 1952, **16**, 269-275.
21. RIESMAN, D. The Lonely Crowd. New Haven: Yale Univ. Press, 1951.
22. STAGNER, R. Homeostasis as a unifying concept in personality theory. *Psychol. Rev.*, 1951, **58**, 5-17.
23. TOLMAN, E. C. Cognitive maps in rats and men. *Psychol. Rev.*, 1948, **55**, 138-208.
24. VERNON, M. D. Visual perception. Cambridge: Cambridge Univ. Press, 1937.
25. WATKINS, J. G., & STAUFFACHER, J. C. An index of pathological thinking in the Rorschach. *J. Proj. Tech.*, 1952, **16**, 276-286.
26. WITKIN, H. A. The nature and importance of individual differences in perception. *J. Pers.*, 1949, **18**, 145-170.
27. ————. Perception of body position and the position of the visual field. *Psychol. Monog.*, 1949, **63**, No. 302. Pp. 46.
28. ————. Personal communication, June 18, 1952.

*Department of Psychology*
*VA Hospital*
*Denver, Colorado*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 151-154

# INTELLIGENCE AND CONSERVATIVE-LIBERAL ATTITUDES*

*Fort Wayne, Indiana*

GRACE RUBIN-RABSON

## A. THE PROBLEM

According to Kerr's (5) summary, a low but significant correlation between liberalism and intelligence is a frequent finding in attitude studies. Two questions arise: (*a*) When the attitude scale produces a total score derived from several sub-sections, is the correlation with intelligence due to a liberalism generality or to some particular sub-section score? (*b*) What specific aspects of intelligence underlie the correlation?

## B. PREVIOUS STUDIES

It might be postulated that information plays some part in attitude formation. Dexter (2) finds a radicalism-general-information correlation of .30; Shimberg (11) a slight but significant difference in attitudes of well and ill-informed, the former group more internationally minded, more aware of the implication of events and less given to the emotional solution of international problems. Nettler (7), unimpressed by the degree of proved dependence, says: "The low relationship of information and prejudice when viewed in company with the empirically demonstrated effectiveness of . . . non-rational motivants of attitudes . . . should make it apparent to the American educator and social meliorist . . . that the academic ideology of fact 'freeing' belief is only partially true."

If information grows during college years, the observation that seniors are more liberal, as Moore and Garrison (6) and Boldt and Shoud (1) concluded, is not unexpected. However, Jones (4) found them less so. Hunter (3) observes that many women students do learn in the course of their college years to go beyond facts and at least initiate the process of thinking.

Kerr (5), reviewing attitude research, finds that information and schooling do for the most part play a rôle; liberal college students tend to possess more relevant information and to do more reading than do conservative students; political-economic liberalism tends to increase with formal schooling; liberal teachers tend to possess more information on current questions.

---

*Received in the Editorial Office on September 19, 1953, and published immediately at Provincetown, Massachusetts. Copyright by the Journal Press.

## C. Procedure

For 49 adults in a psychology class at Indiana University in Fort Wayne, ages 18 to 58 years, intelligence quotients yielded by the verbal section of the Bellevue-Wechsler Intelligence Scale (except digit-memory) are available, as well as total and sub-section scores of Hunter's Test of Social Attitudes. The latter comprises seven sections concerned with the Negro, war, economics and labor, social life and convention, government, religion and "miscellaneous."

According to Hunter's norms, the present group was exceeded in liberalism by only 40 per cent of the population. An examination of the sub-scores, however, reveals that the liberal Negro attitude score, exceeded by only 20 per cent of the population, accounts for this result. Without this score, the group becomes highly conservative, with 60 per cent of the population more liberal.[1]

## D. Results

Table 1 presents the means and standard deviations of the Bellevue-Wechsler verbal *IQ*[2] and of the weighted scores of the five sub-tests, as well as the correlations for the attitude score with and without the Negro score.

TABLE 1

| | Total Att. score | Total minus Negro score | Mean | SD |
|---|---|---|---|---|
| B-W verbal | .40** | .19 | 114.14 | 9.43 |
| Information | .34* | .16 | 11.46 | 2.27 |
| Comprehension | .13 | −.07 | 11.36 | 1.80 |
| Arithmetic | .29* | .14 | 12.12 | 2.71 |
| Similarities | .33* | .24 | 11.88 | 2.34 |
| Vocabulary | .51** | .25 | 11.28 | 2.25 |
| B-W and Negro score | .38** | | | |

*Significant at the 5 per cent level.
**Significant at the 1 per cent level.

It will be seen that the five significant correlations disappear when the Negro score is removed. The reason for this lies presumably in the correlation of .38 between intelligence and liberalism in the Negro score, significant at the 1 per cent level. It is clear too that the five sub-tests influence the correlation in various degrees. "Comprehension" might logically operate in any test, yet here it fails to play a significant rôle. Again, within the qualification that the correlation vitality inheres primarily in the Negro attitude score, information and some competence in handling concepts are also active in effecting a relationship.

---

[1] Analyses of group attitudes and of the Negro score appear elsewhere (8, 9).

[2] The high mean is somewhat belied by the Otis mean of 105 for the same subjects. For a further discussion of this point, see (8).

The participation of arithmetic, whatever of true reasoning may be tapped in the test, is the more interesting since for this group (*IQ* range: 98-138) the first seven items of the test were non-discriminating (10); hence the correlation rests on the variations among Items 8, 9, and 10.

Whatever the validity of the sub-tests, vocabulary enjoys general acceptance as the single most discriminating measure of intelligence. The correlation of .51, significant at the 1 per cent level, lends some corroboration to the conclusion that an attitudes-intelligence relationship may indeed exist.

## E. SUMMARY AND CONCLUSIONS

For 49 adults, total scores for Hunter's Test of Social Attitudes and intelligence quotients produced by the verbal section of the Bellevue-Wechsler Intelligence Scale show a low but significant correlation between liberalism and intelligence. However, when the aberrant liberal Negro score is removed, this relationship disappears.

With the exception of "comprehension," the several sub-tests of the verbal section of the Bellevue-Wechsler also reveal low but significant relationships with total attitude score. The fairly high "*r*" of vocabulary and attitudes corroborates the conclusion that an intelligence-liberalism relationship does indeed exist. But without the Negro score, these correlations too lose vitality.

Why intelligence seems to operate only in the area of Negro attitudes is not clear. That it does so, however, suggests caution in interpreting results of scales which summarize attitudes toward more than one institution or area under an overall score.

## REFERENCES

1. BOLDT, W. J., & SHOUD, J. B. Changes in the attitudes of college students. *J. Educ. Psychol.*, 1934, **25**, 611-619.
2. DEXTER, E. Personality traits related to conservatism and radicalism. *Charac. & Personal.*, 1938, **7**, 230-237.
3. HUNTER, E. C. Changes in general attitudes of women students during four years in college. *J. Soc. Psychol.*, 1942, **16**, 243-257.
4. JONES, E. S. Opinions of college students. *J. Appl. Psychol.*, 1926, **10**, 427-416.
5. KERR, W. A. Correlates of politico-economic liberalism-conservatism. *J. Soc. Psychol.*, 1944, **20**, 61-77.
6. MOORE, G., & GARRISON, K. C. A comparative study of social and political attitudes of college students. *J. Abn. & Soc. Psychol.*, 1932, **27**, 195-208.
7. NETTLER, G. Relationship between attitude and information concerning Japanese in America. *Amer. Sociol. Rev.*, 1946, **11**, 177-191.

8. RUBIN-RABSON, G. Several correlates of a conservatism-liberalism attitudes scale. *J. Soc. Psychol.*, (in press).

9. ———. Liberalism toward Negroes as a deviant reaction in a conservative group. *J. Soc. Psychol.*, (in press).

10. ———. Item order in four verbal sub-tests of the Bellevue-Wechsler Scale. (unpublished).

11. SHIMBERG, B. Information and attitudes toward world affairs. *J. Educ. Psychol.*, 1949, **40**, 206-222.

*1001 East Rudisill Blvd.*
*Fort Wayne, Indiana*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 155-162.

# SOCIAL PERCEPTION AS A FUNCTION OF IDENTIFICATION*

*Galesburg State Research Hospital, Galesburg, Illinois*

LEO SCHNEIDERMAN

## A. PERCEPTION

It has been suggested by Gibson that a theory of schematic perception might start out by predicting ". . . that in two situations identical as regards external stimulation but different as regards motivation, the resulting schematic percepts will be different" (4, p. 105). If one were to apply this observation to the problem of how people perceive each other (2, 7, 11, 14), the expectation might be entertained that differences in set could give rise to different schematic perceptions of other people (6). A further step would be the assertion that *preferred attitudes,* consistent with the identity of the perceiver, may influence his experience (8). If it is assumed that the identity of the perceiver can be defined in large part by the relative consistency of his behavior, it becomes possible to ask if that same consistency does not make for perceptual constancy, as the process mediating between ego structure and the executive act. The possible adaptive significance of constancy behavior, in this connection, has been noted by Ittelson (5), who has called attention to the contribution of the perceiver's accumulated experience in maintaining a stable subjective world. It is not necessary to suppose that perceptual constancy, reflecting the individual's preferred attitudes, operates so as to override the contribution of stimulus determinants. It is enough, for our purpose, to acknowledge that ". . . stimuli cannot be described without reference to the organism . . ." (1, p. 311). But what if the stimuli in question are other people? Does it not become necessary, in such cases, to take into account the relationship between the identity of the object and that of the perceiver? The definition of the stimulus now becomes a matter of referring to *two* organisms, and suggests that the stimulus properties of a person *qua* object may reflect the nature of his relationship to the perceiver.

Freud has invoked the operation of repression as an adaptation to threat originating within the individual himself. He has argued, in this regard, that in obsessional neuroses, a "hostile impulse against a loved person . . . has

---

*Received in the Editorial Office on October 5, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

undergone repression" (4, p. 95). It is clear that repression here denotes the avoidance of a consciously painful experience that would result from acknowledging, or acting upon, some impulse in relation to another person. That the other person is a *loved* person is stated, but not enlarged upon. Whether or not one agrees with Freud's analysis of what happens in obsessional neuroses, his formulation at least calls attention to the particular way one feels toward the other person as a possible determinant of one's perception of him.

Since people differ conspicuously in the rôles they assume, it is possible that perception of the other person may be influenced by the perceiver's feeling of relatedness to him within a broad context of social expectations, as well as by personal feelings. For example, one never perceives an angry face, but rather the angry face of someone who is identifiable with respect to age, sex, race, status, degree of familiarity, etc., and toward whom one stands in some presumed relation, as a consequence. Even one's own diffusely perceived emotional states, such as vague discontentments, or apprehensions, are related probably to past outcomes and present anticipations of one's relationships to others. A perceptual analysis of social interaction might start, then, with a consideration of how the perceiver sees himself in relation to other persons within a given situation.

It might be advantageous to consider leadership behavior in small groups (13), for example, in terms of how each group member perceives himself in relation to other members, and to suggest some hypotheses as to why certain persons are perceived as leaders.

The perceptual field of each group member might be described as including a number of persons, each of whom stands in some relation to the perceiver, and to the other members of the group. The extent to which any member will "stand out" from the others, that is, serve as *figure* against the background of the group, will be influenced probably by many factors, including, perhaps, each member's conception of the degree of homogeneity of the group. Homogeneity is here used in the sense of denoting those perceived, or assumed traits of group members that lead a given participant to think he has something in common with the others. As MacLeod puts it, ". . . within the category of psychologically apprehended persons there is a range of variation which extends from persons who are relatively neutral or indifferent to us to persons who are highly charged with selfhood . . ." (10, p. 203). Since each member is in the position of having to attempt some degree of differentiation of his field, his relative ability to break up the field into figure and ground might determine his success in focusing upon individual members of the group, or even upon himself as a contributor to the group It might be predicted, therefore, that if all the members conceive of themselves as belonging to a homogeneous group, individual differences between members will not be perceived clearly. Also, each member will be unable to use his own feelings, attitudes, values, i.e., his self-related standards, as a point of reference for evaluating what is happening to him. Although this conclusion is in apparent contradiction with the point of view that people are better judges of those like themselves (12), we are dealing here with *presumed* similarity, without reference to the validity of the analogy.

In Lewinian terms, the situation might be described as one of weakly defined boundaries between individuals, with the possibility of distinguishing between self and non-self greatly reduced. In such a situation, it would be impossible for a leader to emerge, unless he could extricate himself, somehow, from the group, and so, bring himself to the attention of the others. The consequence for the "leader" of differentiating himself from the others might be that he would come to be perceived as "different" from them. A polarization might occur, in which the leader would find himself sealed off from the rest, and in their view, no longer sharing in the common identity of the group.

Here is a paradox! How can one be a leader, and yet, be perceived as a stranger? The Lewinian analogy of psychological boundaries may be of some help here. In a group that is judged by each member to be homogeneous, it is possible to speak of each participant as if he had more or less equal access to each of the others. That is, a move toward interaction with another would not involve the risk of being rebuffed, since, in effect, there would be no attempt to cross a clearly defined boundary. Using Stagner's (15) homeostatic analogy, one might speak of a situation in which each perceiver sees himself in relation to an environment that is in a constant state. To maintain equilibrium between himself and the others in his group, each member submerges his own identity, and does not focus on the discrete identities of the others. As a result, each member interacts with the others in an unselfconscious way. The emergence—by whatever means—of a single individual to a position of prominence might be accompanied by a quickening of each member's sense of "apartness" from him, such that the boundaries surrounding the latter would become relatively well defined. To understand how the "prominent" individual can become a leader, it is necessary to consider how his behavior comes to be positively valued, despite the fact of his apartness from the group.

Werner (17) has called attention to physiognomic perception as a type of experience in which the expressive qualities of a stimulus are perceived directly, without distinguishing them from other aspects of the object. Although it is possible to think of physiognomic qualities as inhering in certain objects, it is clear that the perceiver must bring to his experience feelings of his own that *motivate* him to focus on certain stimulus properties, and not on others. For example, the "threatening" quality of a person's appearance is more or less frightening, depending on whether one is afraid of him, or of people generally, to begin with. Similarly, the "friendly" quality of a person's appearance is attractive, or unattractive, depending on whether one is in a bitter, suspicious frame of mind, or at peace with the world. With respect to leadership behavior, it is obvious that each person's expressive qualities, as well as the content of his actions, must have an effect on the remaining members of the group. Even in a group that is thought to be homogeneous by its members, certain individuals will appear in a more conspicuous light than others, depending on their participation, or nonparticipation at certain points in the group's transactions. Also, the ways in which certain individuals acknowledge or ignore the contributions of others, the extent to which they appear to be complacent, determined, cynical, aroused, controlled, confused, defensive, amused, etc., will distinguish them from other group members. If all other things were equal, the leader might be that person whose behavior, by combining certain expressive and other qualities, rendered him most "popular," depending on the predilections of the group, i.e., their evaluation of his behavior in the light of their feelings toward him.

## B. IDENTIFICATION

If this is so, one may take the position that it is possible for two or more persons in a group to show similar behavioral characteristics, and yet, only one will be perceived as a leader. Two persons may be equally aggressive and negativistic, for example, in their behavior toward other group members, yet one will be detested generally, and other acknowledged as a leader. This possibility is important for our argument because it suggests the necessity for defining the leader-follower relationship in terms not only of the stimulus value of the leader, but also of the follower's readiness to accept or reject certain of the latter's actions. It is our contention that the kind and degree of identification that one has with another person will determine one's readiness to accept or reject his actions. If one identifies positively with another person, that is, one feels, even upon minimal acquaintance, that there exists an important mutual bond between oneself and the other person, the perception of his actions will tend to reflect one's positive feelings in a physiognomic way.

If the perceiver feels buoyant, he perceives the other person as exalted also; if the perceiver is angry, he perceives the expression of the other person as angry also. But if the perceiver does not identify with the other person, that is, considers him to be "different" in a negative sense, or if he imagines that someone else might perceive the other person as deviant in some way, then the actions of the latter, no matter how appropriate to the situation, are seen as unrelated to the perceiver's feelings, or even, as opposed to them.

For a group member to function as a leader, then, it is necessary that his actions be imbued with positive feelings by the others. Thus, if the leader expresses himself in an intense, vehement manner, he is perceived as "forceful." If another group member, with whom the others feel only a slight sense of kinship, expresses himself with like intensity, he is perceived as "hot headed," or "indiscreet." In the latter instance, the perceiver has selected out his own negative, or indifferent feelings toward the other person, and has projected them onto the latter's actions. The behavioral configurations presented by leader and nonleader respectively, may be objectively alike, yet be perceived in different ways, as has been suggested. Leader and nonleader, in their behavior, may differ conspicuously, and in the same direction, from the other group members, yet the behavior of the leader will be positively valued, and that of the nonleader negatively valued. Returning to the Lewinian model, it is possible to describe the boundary between leader and group as being just as clearly defined as it is between, let us say, a rejected person, and the rest of the group. The type of transaction between individual and group is similar in both instances: the members of the group perceive the behavior of the individual as "standing out" against the background of their feelings toward him. The individual, in turn, depends on the group members to define his rôle in terms of their feelings. The leader, however, has access to the others by virtue of the fact that they are identified with him. As a result, he is in a position to influence group decisions. The rejected person lacks access to the other group members to the extent that they fail to identify with him. Presumably, any individual with whom the other group members are identified, can function as a leader, providing, also, he can show certain personal qualities of leadership, whatever they may happen to be.

Selective perception by group members operates, then, to insure a favorable reception to the actions of that individual with whom the others have made a positive identification. However, the members must do more than perceive the "good" qualities of the person they like: they must ignore his "bad" qualities. The group members respond to their leader very much in the manner of a person of subordinate status who does not resent his superiors, even though they are oppressive. Selective perception serves here not only as a defense against the perception of threat, but also as a means by which the perceiver, by identifying himself with a valued object, may define his own identity in the most flattering terms.

Using the concept of identification in its present context, it is possible to understand how large social and cultural aggregations, as well as face to face groups, are led to perceive their environments in a selective way. In particular, the member of a marginal group—the immigrant, the member of a racial minority, the newcomer to a community, the nouveau riche—who is in the position of having formed multiple, inconsistent identifications, is impelled by the urgent demands of his situation, to perceive his environment in a selective way. He must acquire quickly a new, and exclusive value orientation; in this process of adaptation, the attraction to the ingroup may prevent the marginal man from perceiving its characteristics in an objective way. By the same token, he may come to perceive the behavior of his own group members in a critical way. Such selective perception presupposes, of course, a situation in which the marginal man has had an opportunity to identify himself with the ingroup, as a result of its partial acceptance of him. Where the ingroup remains consistently inaccessible, and even hostile, there is a strong likelihood that the marginal man will be unable to identify himself with its members, and will perceive their behavior in keeping with his feelings of alienation. For example, the immigrant generation of most ethnic groups in the U. S. is distinguished often by its reluctance to undergo acculturation, in contrast to the second generation, which often accepts acculturation eagerly (16). Obviously, the separatism of the first generation is not the result of choice entirely, since the larger society is prone to define for the immigrant a circumscribed rôle in its economic and social life. As Lewin (9) has pointed out, the leaders of the marginal group, those who have entered with some success into the activities of the ingroup, characteristically are divided in their loyalties, and often may favor the interests of the larger collectivity over those of their own group.

The consistency, or inconsistency of the ingroup, in its treatment of the marginal group, resembles, in its effects, the actions of a leader toward the members of his group. The leader who consistently maintains a barrier between himself and the others through negativism, indifference, abusiveness, etc., ceases to be a leader because he makes it impossible for anyone to continue to identify with him. Only where there is present a pathological degree of self hatred in the members of a group, will the consistently frustrating leader find a following whose expectations are in accord with his actions. Thus, the differentiation of the leader from the rest of his group must not preclude friendly transactions between the former, and at least some of the other group members. Otherwise, the leader will cease to be perceived against the background of the group's positive identification with him, and will stand out merely as "different," on the basis of his expressive, and other personal characteristics.

## C. SUMMARY

Attention has been drawn to the necessity for treating social perception in terms of actual interaction between perceivers. A few illustrations have been provided for the suggestion that people perceive each other in terms of their general attitudes, including their feelings of identification with one another. An attempt has been made, in this connection, to define leadership choice in small groups by reference to the constructs of identification and selective perception. Implicit in our discussion is the notion that each person maintains perceptual constancies in the face of variegated social experiences. That is, depending on one's general disposition toward another person, one will perceive objectively unrelated actions and expressions on the part of the latter as fitting into a consistent pattern of behavior. The actions of a person with whom one is positively identified will be perceived physiognomically, with the perceiver transforming the expressive qualities of the object by the projection of his own positive feelings. Similarly, the actions of a group toward whom one feels psychologically distant, will be perceived physiognomically, and in a constant way, in keeping with one's attitude of estrangement. It is suggested that an intensional approach to perception in terms of personality, must deal with the possibility that equivalent responses reflect functional unities in personality organization, as well as in the configurations of external objects.

## REFERENCES

1. Bruner, J. S. One kind of perception: a reply to Professor Luchins. *Psychol. Rev.*, 1951, **58**, 306-312.
2. Estes, G. S. Judging personality from expressive behavior. *J. Abn. & Soc. Psychol.*, 1938, **33**, 217-236.
3. Freud, S. Collected Papers. (Vol. 4.) London: Hogarth, 1950.
4. Gibson, J. J. Theories of perception. In Dennis, W. (*Ed.*), *Current Trends in Psychological Theory*. Pittsburgh: Univ. Pittsburgh Press, 1951. (Pp. 85-110.)
5. Ittelson, W. H. Constancies in perceptual theory. *Psychol. Rev.*, 1951, **58**, 285-294.

6. Jacobs, D. F., & Pearl, D. The relationship of sociometric choice pattern to visual perception thresholds. *Amer. Psychol.,* 1953, **8**, 373 (Abstract).
7. Kanner, L. Judging emotions from facial expressions. *Psychol. Monog.,* 1931, **41**, No. 186, 1-91.
8. Klein, G. S., Schlesinger, H. J., & Meister, D. E. The effect of personal values on perception. *Psychol. Rev.,* 1951, **58**, 96-112.
9. Lewin, K. Resolving Social Conflicts. New York: Harper, 1948.
10. MacLeod, R. B. The phenomenological approach to social psychology. *Psychol. Rev.,* 1947, **54**, 193-210.
11. McTeer, W. Observational definitions of emotion. *Psychol. Rev.,* 1953, **60**, 172-180.
12. Notcutt, B., & Silva, A. L. M. Knowledge of other people. *J. Abn. & Soc. Psychol.,* 1951, **46**, 30-37.
13. Roseborough, M. E. Experimental studies of small groups. *Psychol. Bull.,* 1953, **50**, 275-303.
14. Scodel, A., & Mussen, P. Social perceptions of authoritarians and nonauthoritarians. *J. Abn. & Soc. Psychol.,* 1953, **48**, 181-184.
15. Stagner, R. Homeostasis as a unifying concept in personality theory. *Psychol. Rev.,* 1951, **58**, 5-17.
16. Warner, W. L., & Srole, L. The Social Systems of American Ethnic Groups. New Haven: Yale Univ. Press, 1945.
17. Werner, H. Comparative Psychology of Mental Development. New York: Harper, 1940.

*1470 North West Street*
*Galesburg, Illinois*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 163-171.

# PERFORMANCE ON A FOLLOWING TRACKING TASK (MODIFIED *SAM* TWO-HAND COORDINATION TEST) AS A FUNCTION OF RADIUS OF CONTROL CRANKS*[1]

*Department of Psychology, University of Rochester*

PAUL SWARTZ, EUGENIA B. NORRIS, AND S. D. S. SPRAGG

## A. INTRODUCTION

This is one of a series of experimental studies of continuous movement reactions as found in tracking tasks. The first investigation dealt with the problem of complex two-hand coördination as a function of the planes of rotation of the control cranks (4). The second study examined two-hand coördination as a function of the relations between direction of rotation of control cranks and direction of movement of display (5). The present experiment was designed to investigate the adequacy of two-hand coördination as a function of the radius of the control cranks.

Previous work by Reed (6) and by Helson (2) has shown that size of controls constitutes an important aspect of equipment design. Reed, using an apparatus which permitted variation in crank radius, braking force, position and orientation of the crank, measured maximal rate of turning under various conditions. Four of his major findings which bear on the present study are the following:

> 1. As radius was varied from 1.5 to 24.0 cm., the angular velocity was maximum at about 2 cm. and reached a minimum at the largest radius. The point at which the maximum occurred was a function of the degree of coordination.
> 2. As radius increased the linear velocity of the hand increased.
> 3. An increase in force required to turn the handwheel affected short radii to a greater extent than large radii.
> 4. Cranking motion in a vertical plane parallel to the front of the body yielded the highest rate of rotation. However, the introduction of a braking force rendered this position "unstable," in which case the preferred plane of rotation was vertical but perpendicular to the front of the body.

---

*Received in the Editorial Office on October 20, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was conducted as part of a research program on human motor performance carried out under Contract No. N6onr-241, T. O. 6 between the University of Rochester and the Special Devices Center, Office of Naval Research, to whom these results have been reported in mimeographed form as Technical Report No. SDC-241-6-6.

Helson's study (2) investigated tracking performance on a type of pursuit task, in which the subject was required to align a moving indicator with a stationary target by means of a handwheel. Among other things he found that performance with a 4.5 inch radius handwheel was consistently better than performance with a 2.25 inch radius wheel up to 100 r.p.m., beyond which point the latter yielded superior performance. The "breaking point" for speed was reached sooner by large than by small handwheels. Small light handwheels yielded superior performance when tracking courses involved rapid reversals in direction and high acceleration. Conversely, tracking tasks which did not exceed a subject's ability to handle motion and loads were best performed with larger and heavier handles.

However, the findings cited above have a limited applicability, since only one-handed cranking was involved. There are many control tasks which require the coördinated cranking of two handles. Little is known concerning the relative adequacy of various control sizes in such situations. The present study is intended to supply information in this area.

## B. Statement of the Problem

The purpose of the present study was to investigate adequacy of performance in a two-handed tracking task (using the modified *SAM* Two-Hand Coördination Test) as a function of the radius of the control handles. Five different control radii, each employed in two operating positions, were used.

## C. Subjects

Forty male undergraduates at the University of Rochester, ranging in age from 17 to 22 years, served as subjects. Both right and left handed subjects were used. They were paid for their services. None of the subjects had had any previous experience with the Two-Hand Coördination Test. To randomize the groupings subjects were assigned to the various experimental conditions in the order of their appearance.

## D. Apparatus

The apparatus employed was a modified *SAM* Two-Hand Coördination Test (CM101B), described in the U. S. Air Force publication, *Standard Classification Tests* (3). This apparatus was designed to measure the ability to coördinate simultaneous rotary movement of the two hands involved in superimposing a target follower on a visually perceived target moving at varying speeds in an irregular path in the horizontal plane. The modifications, which have been fully described in a previous report (4), consisted basically in mounting adjustable handles on pivoting stands, thus making it possible to vary the radius of the controls and place them in either the horizontal or the vertical plane. A paired synchro drive system was used.

Each adjustable handle consisted of a 9-inch crank arm mounted on a 3/8 inch steel shaft through a keyed shaft hole three inches from one end of the arm. A 5.5 inch milled slot along the long section of the arm permitted the positioning of the crank handle within a range of 1.0 to 5.5 inches from the center of the shaft. A 2.25 inch slot along the short section of the arm permitted the positioning of a balancing disc for all crank radii within the above range. The arm, handle, and balancing disc were all constructed of aluminum and together weighed 1.25 lbs.

## E. PROCEDURE

Each subject served in one experimental session, at the beginning of which the general instructions devised for a previous study (4) were read.

Each subject was then given two trials on each of the following control radii: 1.0, 2.0, 3.0, 4.0, and 5.5 inches. The trials for each control radius were run successively, making five blocks of two trials each. A latin square design was employed to balance the sequence of presentation of these experimental conditions. There were four replications of the design for each position. One trial consisted of one complete revolution of the target, lasting one minute. Each block of trials was separated from the preceding one by a one-minute rest interval. Within a block, 15 seconds intervened between trials.

Twenty subjects performed in Position I, in which both control cranks rotated in a vertical plane parallel to the frontal body plane. Twenty performed in Position IV, in which the right control crank rotated in a vertical plane perpendicular to the frontal body plane and the left control crank was placed as in Position I. [The designations "Position I" and "Position IV" were those originally assigned to these two among the several combinations of planes of rotation described in the first paper in this series (4).] For both positions clockwise rotation of the right control crank moved the target follower directly away from the operator while clockwise rotation of the left control crank moved the target follower to the right.

Scoring (time on target) was achieved by means of a Standard Electric Timer, calibrated in thousandths of a minute.

## F. RESULTS

The results have been expressed in terms of mean time on target for each of the control radii within each position and for Position I vs. Position IV. In Table 1 are presented the group means in milliminutes for each control radius for Positions I and IV respectively. Figure 1 shows the same data graphically.

TABLE 1

MEAN TIME ON TARGET (IN MILLIMINUTES)

| Control radius (in inches) | Position I | Position IV |
|---|---|---|
| 1 | 444.0 | 506.4 |
| 2 | 549.9 | 610.7 |
| 3 | 570.3 | 615.5 |
| 4 | 519.6 | 602.1 |
| 5.5 | 494.4 | 572.1 |

TABLE 2

ANALYSIS OF VARIANCE FOR GROUPS I AND II, AND FOR GROUP I VERSUS GROUP II

| Source | df | MS | F ratio | Sig. Level |
|---|---|---|---|---|
| Position I (Group I) | | | | |
| 1. Size | 4 | 48,749.1 | 7.90 (1/5) | .001 |
| 2. Sequence of size | 4 | 32,656.6 | <1 (2/4) | — |
| 3. Ordinal position of size in sequence | 4 | 307,791.5 | 49.91 (3/5) | .001 |
| 4. Individual subject in sequence | 15 | 59,990.0 | 9.73 (4/5) | .001 |
| 5. Error (combined term: error & sq. uniqueness) | 72 | 6,167.15 | | |
| 6. Total | 99 | | | |
| Position IV (Group II) | | | | |
| 1. Size | 4 | 41,051.4 | 10.42 (1/5) | .001 |
| 2. Sequence of size | 4 | 63,111.1 | <1 (2/4) | — |
| 3. Ordinal position of size in sequence | 4 | 362,507.8 | 92.00 (3/5) | .001 |
| 4. Individual subject in sequence | 15 | 167,943.3 | 42.62 (4/5) | .001 |
| 5. Error (combined term: error & sq. uniqueness) | 72 | 3,940.27 | | |
| 6. Total | 99 | | | |
| Position I vs. Position IV | | | | |
| 1. Groups | 1 | 215,955.96 | 1.89 (1/2) | NS |
| 2. Individual subjects | 30 | 113,966.7 | | |
| 3. Error (combined term: error & sq. uniqueness) | 144 | 5,053.71 | | |

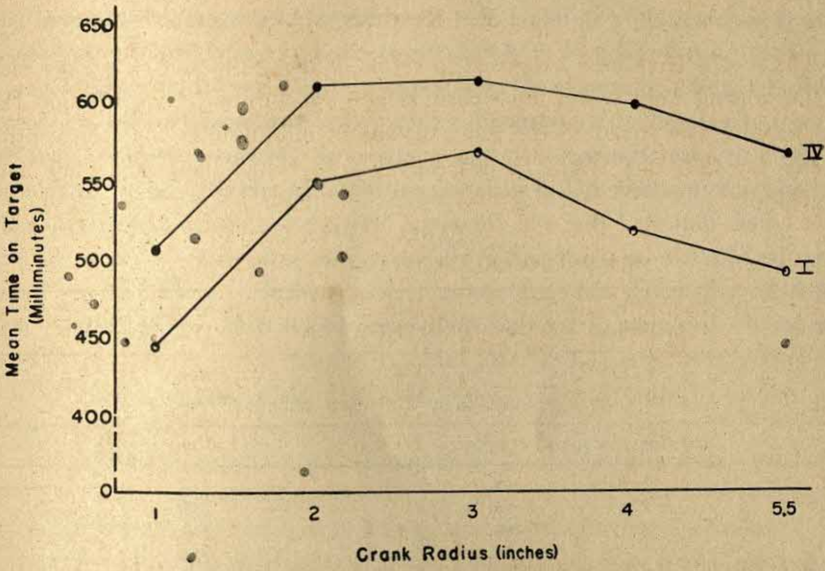


FIGURE 1
SHOWING THE MEAN TIME ON TARGET AS A FUNCTION OF CONTROL CRANK RADIUS, FOR POSITION I (GROUP I) AND FOR POSITION IV (GROUP II)

In Table 2 the analyses of variance are summarized for Position I and Position IV separately, and for Position I vs. Position IV. The procedure followed was that described by Grant (1), allowing a pure error-estimate of variation which is independent of sizes, sequence of sizes, ordinal position of size within sequence, and the individual differences among the subjects within a sequence. Since, for both positions, a preliminary analysis showed that the $F$ ratio for square uniqueness over the error term was less than 1.0 the error term was combined with the square uniqueness term. For both positions sizes, sequence of sizes, and the ordinal position of size within sequence are all significant beyond the .001 level when tested against the error estimate. The lack of significance of the $F$ ratio for sequences over individual subjects within sequences indicates that the significant $F$ for sequences over the error term can be accounted for in terms of significant variation between subjects. The significance of the mean square for individual subjects within sequence attests to the reliability of the scores used.

Since the analysis of variance indicated that within Position I the variance due to crank radius is significant beyond the .001 level of confidence the $t$-test was applied (using the error variance of the analysis of variance as the error estimate). It was found that the difference between the means of any two radii must be 65.8 to be significant at the 1 per cent level of confidence, 59.1 at the 2 per cent level, and 49.66 at the 5 per cent level. Table 3 presents the levels of confidence for those differences found to be significant.

Similarly, for Position IV the analysis of variance revealed that the variance due to crank radius is significant beyond the .001 level of confidence. The *t*-tests indicated that the difference between the means of any two radii must be 52.60 to be significant at the 1 per cent level of confidence, 39.70 at the 5 per cent level, and 38.31 at the 6 per cent level. In Table 4 are found the levels of confidence for those differences found to be significant.

TABLE 3

SIGNIFICANCE OF DIFFERENCES BETWEEN MEANS: POSITION I

| Radius of control cranks (in inches) | | | Level of confidence |
|---|---|---|---|
| 2, 3, 4 | vs. | 1 | 1% |
| 5.5 | vs. | 1 | 5% |
| 2 | vs. | 5.5 | 5% |
| 3 | vs. | 5.5 | 1% |
| 3 | vs. | 4 | 5% |

TABLE 4

SIGNIFICANCE OF DIFFERENCES BETWEEN MEANS: POSITION IV

| Radius of control cranks (in inches) | | | Level of confidence |
|---|---|---|---|
| 2, 3, 4, 5.5 | vs. | 1 | 1% |
| 2 | vs. | 5.5 | 6% |
| 3 | vs. | 5.5 | 5% |

It will be noted that for both positions the 2, 3, 4 and 5.5 inch controls all yielded significantly better performance than the one-inch control. Also, for both positions the 2- and 3-inch controls both yielded significantly better performance than the 5.5 inch control. In addition, for Position I, performance was significantly better on the 3-inch control than on the 4-inch one. Thus, for Position I, the "optimal band" of control radii includes only the 2- and 3-inch controls, while for Position IV, the band is wider, containing the 2, 3, and 4-inch controls.

The analysis of the differences in performance between Position I and Position IV constitutes a "between groups"-"within groups" analysis (see Table 2). The mean square for positions (groups) was divided by the pooled mean square for subjects within sequence. This ratio was not significant. This would indicate that the differences between positions can be accounted for in terms of reliable differences between subjects. The fact that performance in Position IV was somewhat superior to that in Position I for every crank radius used may prove to be significant but this conclusion is not supported by the results of the present experiment. However, the obtained superiority of Position IV agrees with the results of other studies which we have carried out, and it is believed that some confidence may be placed in the finding.

## G. DISCUSSION

The results reported here show that performance on a complex, two-hand coördination task is importantly related to the size (radius) of the control cranks. For both crank position combinations studied, the results show an optimal order of magnitude for crank radius (2 to 3 inches); greater and lesser magnitudes of crank radius yield poorer performance.

These results are consistent with Helson's (2) analysis of tracking behavior, and provide experimental corroboration of his prediction that, "were extremes of size investigated it is very probable that we should obtain a *U* curve for accuracy as a function of size with the sizes employed in this study falling in the optimal band" (2, p. 481).

Comparing the crank radii used by Helson (2.25 and 4.5 inches) with the present results, it is seen that his 2.25 inch crank falls within the "optimal band" for both Position I and Position IV. His 4.5 inch crank does so for Position IV but falls somewhat outside the "optimal band" for Position I.

In the present study the relatively poor performance with the 1- and 5.5-inch cranks in both positions is probably accounted for by the following two performance limiting factors described by Helson:

> First, the course was moderately difficult, requiring sharp reversals of direction and frequent accelerations and decelerations. Since the force required to make these responses is an inverse function of diameter we would expect the shortest control radius to yield the poorest performance
>
> Second, with increase of control radius beyond an optimal value the operator begins to approach his upper limit of speed of cranking. It seems reasonable to hold that when the two-hand coördination task requires simultaneous movement of two 5.5-inch cranks we approach such a condition (2, pp. 486-487).

The results for Position I as opposed to Position IV are in agreement with a previous report from this laboratory (5) and with the findings of Reed (6). For each control radius, performance was somewhat but not statistically significantly better for Position IV than for Position I. This fact might be restated by saying that Positon IV is a less difficult task than Position I.

Our findings also suggest that the optimal range of control sizes may vary inversely with the difficulty of the task to be performed. This relates to the more general hypothesis that the more difficult the task the more important becomes the problem of control design and the use of optimum control values.

## H. SUMMARY

The present study was designed to investigate the adequacy of performance of a two-hand tracking task (using the modified *SAM* Two-Hand Coordination Test) as a function of the radius of the control cranks. Forty subjects were used; each subject was given two trials on each of the following control crank radii: 1.0, 2.0, 3.0, 4.0, and 5.5 inches. A latin square design was employed in order to balance the sequence of presentation of these experimental conditions.

Two combinations of control crank positions were employed. In Position I both control cranks rotated in the vertical plane parallel to the frontal body plane. In Position IV the right-hand control crank rotated in the vertical plane perpendicular to the frontal body plane, and the left-hand crank in the vertical plane parallel to the frontal body plane—i.e., as in Position I. (The designation of these Positions as I and IV is explained in the report.) For both positions clockwise rotation of the right control crank moved the target follower in a direction directly away from the operator; clockwise rotation of the left control crank moved the target follower to the operator's right. Twenty of the subjects performed in Position I, the remaining 20 in Position IV.

The results were analyzed in terms of mean time on target for each of the control radii within each position and for Position I vs. Position IV. Performance was found to be superior for the middle values (2.0, 3.0, and 4.0 inch radii) than for the smallest and largest radii used (1.0 and 5.5 inches). A more detailed analysis showed that both for Positions I and IV the 2.0, 3.0, 4.0, and 5.5 inch control cranks all yielded significantly better performances than the 1.0 inch crank. Also, for both positions, the 2.0 and 3.0 inch cranks were significantly superior to the 5.5 inch crank. In addition, for Position I performance on the 3.0 inch crank was significantly superior to that on the 4.0 inch crank. Finally, for every control radius used performance was somewhat better when the tracking task was performed in Position IV than it was in Position I.

It is concluded that for this and comparable two-hand tracking tasks the optimal magnitude of control crank radius is of the order of two to three inches.

## REFERENCES

1. GRANT, D. A. The latin square principle in the design and analysis of psychological experiments. *Psychol. Bull.*, 1948, **45**, 427-442.
2. HELSON, H. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473-497.
3. MELTON, A. W. (*Ed.*) Apparatus tests. *Report No. 4, U. S. Army Air Forces Aviation Psychology Reports.* Washington, D. C.: U. S. Government Printing Office, 1947. (Pp. 177-259.)
4. NORRIS, EUGENIA B., & SPRAGG, S. D. S. Performance on a following tracking task (Modified *SAM* Two-Hand Coördination Test) as a function of the planes of operation of the controls. *J. of Psychol.*, 1953, **35**, 107-117.
5. NORRIS, EUGENIA B., & SPRAGG, S. D. S. Performance on a following tracking task (Modified *SAM* Two-Hand Coördination Test) as a function of the relations between direction of rotation of controls and direction of movement of display. *J. of Psychol.*, 1953, **35**, 119-129.
6. REED, J. D. Factors influencing rotary performance. *J. of Psychol.*, 1949, **28**, 65-92.

*Department of Psychology*
*University of Rochester*
*Rochester 3, New York*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 173-183.

# TRANSFER EFFECTS BETWEEN PERFORMANCE ON A FOLLOWING TRACKING TASK (MODIFIED *SAM* TWO-HAND COORDINATION TEST) AND A COMPENSATORY TRACKING TASK (MODIFIED *SAM* TWO-HAND PURSUIT TEST)*[1]

*Department of Psychology, University of Rochester*

B. G. Andreas, R. F. Green, and S. D. S. Spragg

## A. Introduction

An analysis of tracking tasks suggests that there are two principal kinds of tracking. The first is "following" tracking in which the operator attempts to keep a "follower" aligned on a "target" with both of these perceived as moving in space. An example is the tracking of an aircraft with an open sight anti-aircraft machine gun. The second type of tracking is designated as "compensatory" tracking in which the operator attempts to adjust controls so as to compensate for the apparent drift in a target, bringing it back to a stationary reference point in the display. An example of a task requiring such compensatory tracking is the operation of a bombsight in which the bombardier tries to keep the target aligned with the cross-hairs of the sight.

Although different, as indicated by the definitions and examples given above, following tracking and compensatory tracking are alike in that both require that adjusting responses, which may be regarded as essentially continuous, be made to any misalignment of reference points in the display. That is, both tasks require the detection and elimination of discrepancy or error. This similarity of the two kinds of tracking task suggests that practice on either may facilitate performance on the other. Other research has found both facilitation and interference when certain aspects of tasks requiring complex coördination have been varied (5, 7, 8, 14). None of these studies so far has used the two different kinds of tracking as the two tasks practiced in succession by the subjects.

---

*Received in the Editorial Office on October 21, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was conducted as part of a research program on human motor performance carried out under Contract N6onr-241, T.O. 6 between the University of Rochester and the Special Devices Center, Office of Naval Research, Port Washington, L. I., N. Y., to whom these results have been reported in Technical Report SDC 241-6-7.

The present study employed the modified *SAM* Two-Hand Coördination Test for following tracking and the modified *SAM* Two-Hand Pursuit Test for the compensatory tracking task. Subjects were given eight trials on one task and then eight trials on the other. Performances with and without prior practice on the other task were compared to determine the transfer effects which might be present.

## B. STATEMENT OF THE PROBLEM

This experiment was designed to determine the transfer effect between a following tracking task and a compensatory tracking task. Four groups of subjects were given eight trials on one task followed by eight trials on the other. Transfer effects were determined by comparing, for each task, performances before and after practice on the other task.

## C. SUBJECTS

The subjects were 92 basic trainees in a military installation. They were assigned to the groups in the order of their appearance. None of the subjects had had previous experience with the apparatus.

## D. APPARATUS

The two pieces of apparatus used were a modified *SAM* Two-Hand Coordination Test (CM101B) and a modified *SAM* Two-Hand Pursuit Test (CM810B) described at length in other reports from this project (11, 12, 13). On both instruments the two control cranks were similarly placed, the left crank rotating in a vertical plane parallel to the frontal body plane of the subject and the right crank rotating in a vertical plane perpendicular to this frontal body plane. These crank positions (which have a continuity of the plane of rotation of each crank with the resultant directions of movement of the controlled display member) had been found to yield best performance on the Two-Hand Coördination Test (12).

## E. PROCEDURE

Four groups of *S*s were used, each *S* being given eight one-minute trials on one task and then eight one-minute trials on the other task. Prior to the first and the ninth trials, *S*s were given standard instructions for the subsequent task. These instructions were essentially similar to those used in air crew testing as cited by Melton (9, p. 192, p. 251).

Compensatory tracking on the Two-Hand Pursuit Test was done with the "natural" control-display relationship for two groups and the "unnatural" for the other two groups. In the "natural" condition the target moved *away from* the subject when the right-hand crank was turned *clockwise,* and vice versa; the target moved *right* when the left-hand crank was turned *clockwise,* and vice versa. This "natural" control-display relationship is illustrated in Figure 1. The "unnatural" condition gives exactly the opposite movement of the controlled member for any direction of crank rotation. For following tracking, only the "natural" condition was employed on the Two-Hand Co-

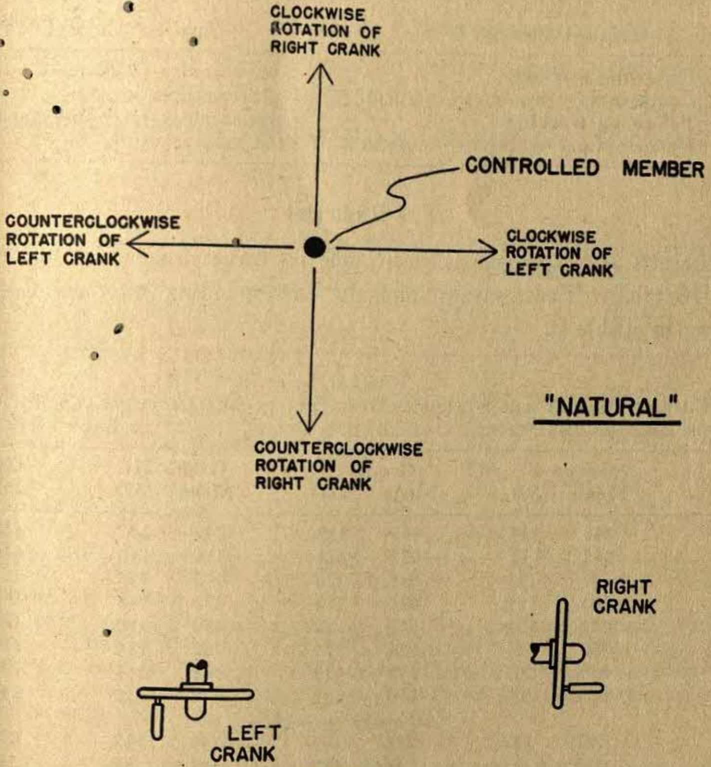

CLOCKWISE ROTATION OF RIGHT CRANK
CONTROLLED MEMBER
COUNTERCLOCKWISE ROTATION OF LEFT CRANK
CLOCKWISE ROTATION OF LEFT CRANK
"NATURAL"
COUNTERCLOCKWISE ROTATION OF RIGHT CRANK
RIGHT CRANK
LEFT CRANK

FIGURE 1

Schematic Top View of Apparatus Showing the "Natural" Control-Display Relationship. The subject's position is at the bottom of the diagram. Controlled Member (Black Dot) is the Target Follower in the Two-Hand Coordination Test, the Target Spot on the Two-Hand Pursuit Test. Each Arrow indicates the Resultant Movement of this Controlled Member when a Particular Crank Rotation is Made. Intermediate Directions of Movement Result When Both Cranks are Turned Simultaneously.

The "Unnatural" Condition Involves Exactly the Opposite Relations Between Direction of Crank Rotation and Movement of the Controlled Member.

ordination Test, the target follower button moving as shown in Figure 1. The measure of performance for either task was time on target for each trial, expressed in milliminutes. The tasks performed by each group are indicated in Table 1.

TABLE 1
TASKS PERFORMED BY EACH GROUP OF SUBJECTS IN EACH BLOCK OF EIGHT TRIALS

| Group | Task "A"<br>Trials 1 through 8 | Task "B"<br>Trials 9 through 16 |
|---|---|---|
| I | Following tracking | Compensatory tracking—"natural" |
| II | Compensatory tracking—"natural" | Following tracking |
| III | Following tracking | Compensatory tracking—"unnatural" |
| IV | Compensatory tracking—"unnatural" | Following tracking |

## F. RESULTS

The results are expressed as mean time on target for each group on each of the 16 trials. These means, and the corresponding standard deviations, are given in Table 2.

TABLE 2
MEAN TIME ON TARGET AND STANDARD DEVIATION IN MILLIMINUTES FOR EACH GROUP FOR EACH OF THE SIXTEEN ONE-MINUTE TRIALS. N=23 IN EACH GROUP.

| Trial | Group I |  | Group II |  | Group III |  | Group IV |  |
|---|---|---|---|---|---|---|---|---|
|  | Mean | *SD* | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| 1 | 240 | 147 | 485 | 219 | 240 | 142 | 542 | 188 |
| 2 | 342 | 171 | 558 | 221 | 336 | 140 | 616 | 183 |
| 3 | 407 | 195 | 591 | 225 | 442 | 136 | 621 | 167 |
| 4 | 401 | 189 | 621 | 210 | 372 | 148 | 645 | 178 |
| 5 | 451 | 184 | 652 | 180 | 492 | 190 | 683 | 175 |
| 6 | 490 | 216 | 626 | 223 | 482 | 156 | 693 | 171 |
| 7 | 574 | 201 | 699 | 193 | 545 | 163 | 727 | 170 |
| 8 | 526 | 206 | 690 | 193 | 548 | 183 | 752 | 170 |
|  |  |  | *Change in task* |  |  |  |  |  |
| 9 | 592 | 182 | 288 | 171 | 758 | 143 | 294 | 170 |
| 10 | 641 | 187 | 351 | 168 | 812 | 86 | 355 | 166 |
| 11 | 684 | 188 | 417 | 172 | 837 | 107 | 389 | 154 |
| 12 | 723 | 169 | 363 | 175 | 831 | 122 | 412 | 191 |
| 13 | 710 | 168 | 427 | 175 | 858 | 94 | 421 | 196 |
| 14 | 735 | 153 | 467 | 199 | 836 | 116 | 475 | 139 |
| 15 | 734 | 122 | 521 | 184 | 878 | 113 | 509 | 187 |
| 16 | 738 | 134 | 481 | 201 | 852 | 97 | 505 | 199 |

Figure 2 represents the performance of Groups I and II on the 16 trials. Group I was given Trials 1-8 on a following tracking task and was then switched to a compensatory tracking task, with the "natural" control-display relationship, for Trials 9-16. The extent to which performance on this second task was influenced by practice on the first task was assessed by comparing the mean score for Group I for Trials 9-16 with that of Group II which performed on the same task on Trials 1-8, without any prior experience in the experiment. The difference was not statistically significant. In other words, the data offer no basis for concluding that Group I's performance on the compensatory tracking task with the "natural" control-display relationship was facilitated by prior practice on the following tracking task.

Table 3 presents the group means for each block of eight trials. The means are paired according to comparability of task and the results of *t*-tests for the differences between these means are given in the last column of the table.

TABLE 3

MEAN TIME ON TARGET SCORES FOR EACH BLOCK OF EIGHT TRIALS ON A PARTICULAR TRACKING TASK, WITH *t* VALUES RESULTING FROM COMPARISON OF PAIRS OF MEANS.

| Type of tracking | Previous task | Group | Mean for 8 trials | *SE of* diff. | *t* |
|---|---|---|---|---|---|
| Following | (none) | I | 429 | 50.0 | .28* |
| Following | Comp.-Nat. | II | 415 | | |
| Following | (none) | III | 432 | 42.2 | .28* |
| Following | Comp.-Unnat. | IV | 420 | | |
| Comp.-Nat. | (none) | II | 616 | 49.2 | 1.60* |
| Comp.-Nat. | Following | I | 695 | | |
| Comp.-Unnat. | (none) | IV | 660 | 34.8 | 4.97** |
| Comp.-Unnat. | Following | III | 833 | | |

*First three $t$ values were not significant; for $t=1.60$ the value of $p$ exceeds .10 for 44 degrees of freedom.

**In computing this $t$ value the variances used were not homogeneous. Following a procedure outlined by Edwards (3, p. 170), the computed value was compared with the tabled value of $t$ for 22 degrees of freedom. Since the tabled value (.001 level) was exceeded by the calculated value, it was concluded that the means, as well as the variances, were significantly different, $p<.001$.

Figure 2 may also be examined for transfer effects from compensatory tracking to following tracking in the performance of Group II. When this group's achievement on the second task (Trials 9-16) is compared with that of Group I which performed initially on the same tracking task (Trials 1-8), there is no statistically significant difference between the means for these blocks of trials as Table 3 shows. Therefore there is no basis for concluding that there is any effect of compensatory tracking practice on subsequent performance on following tracking under the conditions of this experiment.

Figure 3 represents the performance of Groups III and IV. It shows

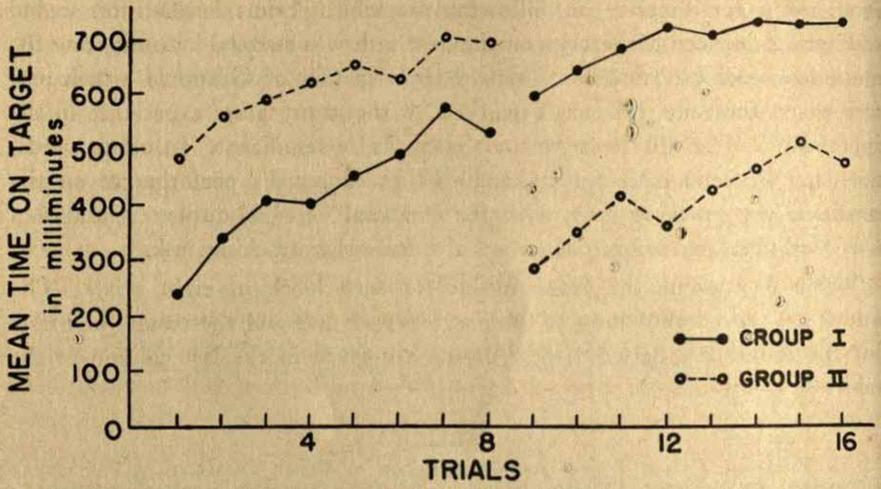


FIGURE 2
Mean Time on Target Scores Made by Group I and Group II
(N=23 in Each Group)

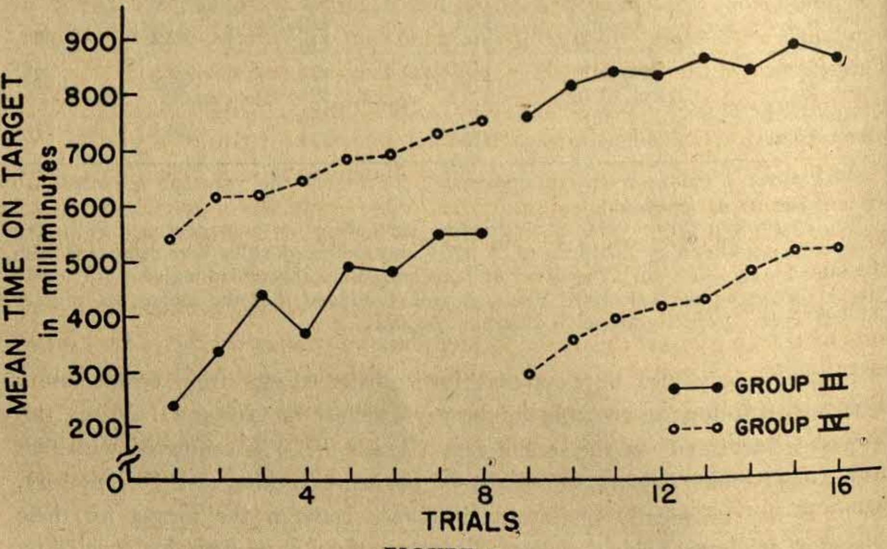


FIGURE 3
Mean Time on Target Scores Made by Group III and Group IV
(N=23 in Each Group)

Group IV's performance on following tracking (Trials 9-16) after eight trials practicing compensatory tracking with the "unnatural" display-control relationship. The achievement of Group IV on Trials 9-16 was compared with that of Group III which worked at the same task during the initial eight trials. Although the curve for Group IV starts slightly higher and ends somewhat lower than that for Group III, the hypothesis that the curves are parallel could not be rejected when a trend analysis (Lindquist, 6, Case 11a, p. 76) yielded an *F* value of only .1. A *t*-test of the grand means of these curves showed them not significantly different, as Table 3 shows. Thus there is no basis for assuming any effect of compensatory tracking practice upon subsequent following tracking performance under the conditions of this study.

Group III was given Trials 1-8 on a following tracking task and was then switched to a compensatory tracking task, with the "unnatural" control-display relationship, for Trials 9-16. This latter task is the same as that performed by Group IV on Trials 1-8 without prior experience in the experiment. A comparison of the performance curves of these two groups (Figure 3) shows that Group III performed better on the compensatory tracking (having previously engaged in eight trials of following tracking) than did Group IV. The two performance curves were compared using a *t*-test and were found to differ by a significant amount as Table 3 shows. This particular result indicates a positive transfer effect or facilitation of performance on compensatory tracking ("unnatural" relationship) resulting from prior practice on following tracking under the conditions of this experiment. Although prior practice on following tracking facilitated subsequent performance on compensatory tracking (Group III), no transfer effect was found when the order of practice of the two tasks was the opposite (Group IV).

In addition to giving information on transfer effects between different sorts of tracking tasks, this study yielded data on compensatory tracking with two different control-display relationships, "natural" and "unnatural." On Trials 1-8 Group IV performed somewhat better with the "unnatural" relationship in effect than did Group II under the "natural" condition. However, a *t*-test applied to the group means for the block of 8 trials showed no significant difference ($t = .88$).

It may be noted, incidentally, that the mean time on target was always greater for compensatory tracking than for following tracking at comparable stages in practice. This indicates that in this experiment the compensatory task was easier than the following task, if time on target is taken as the measure of difficulty. This result was probably due largely to greater target course complexity in the Two-Hand Coördination Test with a resultant greater demand for cranking speed and accelerations.

## G. CONCLUSIONS AND DISCUSSION

The conclusions given below should be considered in the light of the particular conditions of this study. The results obtained were probably dependent to some extent on factors like the similarity of mechanical controls (cranks) used on the tracking devices and the difference in target speeds and courses between the two tasks. Considering the modified Two-Hand Coordination Test and the modified Two-Hand Pursuit Test as representative of following and compensatory tracking, respectively, the following conclusions may be drawn:

1. In many instances, practice on one type of tracking may not facilitate performance on the other.
2. Practice on following tracking does tend to facilitate subsequent compensatory tracking when the "unnatural" control-display relationship is in effect for the latter task.
3. The "unnatural" control-display relationship tends to yield higher scores in compensatory tracking than does the "natural" relationship.

The observation that facilitation effects do not appear generally when subjects are switched from one type of tracking to another is in agreement with findings reported by Brown (1).

It is notable that the "unnatural" control-display relationship appears easier than the "natural" relationship in compensatory tracking. Although the difference obtained in this experiment was not significant, differences observed in other experiments (4, 11) tend to support this conclusion. This notion is also given tangential support by the fact that it was compensatory tracking with the "unnatural" relationship that benefited demonstrably from prior practice on following tracking with the "natural" control-display relationship.

The term "natural" was applied in earlier reports (13) to that control-display relationship which, for following tracking, was in accord with the population stereotype (2, 15), giving the "expected" movement of the display member when controls were manipulated; and which gave higher performance scores. For the present experiment the term "natural" was extended to compensatory tracking upon logical grounds with the target spot (controlled member) of the Two-Hand Pursuit Test considered analogous to the target follower of the coördination test.

In view of the empirical evidence now at hand (relative difficulty and transfer effects) it is the "unnatural" relationship for compensatory tracking which appears functionally equivalent to the "natural" relationship for following tracking. A certain equivalence of formal definition would be achieved if, for compensatory tracking, the *apparent* or *relative* motion of the *cross-hairs* were considered the display referent in defining the control-display relationships. This change in referent would interchange the "natural" and "unnatural" designations for compensatory tracking from those used in this report. With this change in terms the "natural" relationship would be the easier for both types of tracking and the single instance of positive transfer observed in this study would be from "natural" following tracking to "natural" compensatory tracking.

It should be noted that the results which showed the "unnatural" control-display relationship to lead to better performance than the "natural" condition all came from studies performed with the modified Two-Hand Pursuit Test (4, 11). Differences found were not always significant nor did they favor the "unnatural" condition in every single instance, although the weight of evidence supported this conclusion. In other studies of compensatory tracking (10) it has been found that the "natural" control-display relationship results in higher scores. These experiments involved both control and display characteristics which were markedly different from those of the present study, a fact which might account for the lack of agreement in the results. In considering control-display relationships it is necessary to consider the *kind* of display and the *kind* of controls being used.

The finding of a net positive transfer effect in only one out of four experimental groups in this study might be accounted for by numerous *ad hoc* explanations based on similarities and differences between the two tracking tasks. Similarities which might have given rise to positive transfer effects might include (*a*) functional similarity of tasks (as reflected in the proposed redefinitions of "natural" and "unnatural" compensatory tracking) of control-display relationships for Groups III and IV, (*b*) general orientation to the two tasks as elicited by instructions, and (*c*) identical controls on both devices.

Negative transfer effects might have arisen from (*a*) functional opposition of the redefined control-display relationships in Groups I and II, (*b*) differences in the display in the two tasks, (*c*) differences in speed and accelerations of cranking required, and (*d*) the change from an easier to a more difficult task in Groups II and IV. These possible negative factors might be variously combined with the positive factors listed above to account for the results obtained. However, since the data offer no basis for a choice among possible explanations, the attempt to outline one seems unwarranted at this time. It might be noted that the change from an easier to a more difficult task in Group IV might have prevented the appearance of a positive transfer effect such as that observed in Group III where the same tasks were performed in the opposite order. This unidirectionality of the positive transfer effect points up the dependence of transfer in psychomotor tasks upon more than a single factor. The fact that positive transfer was not observed in Group I which also transferred from a relatively more difficult to an easier task might, for instance, mean that the potential positive effects of this variable were counterbalanced by the potential negative effects of transferring from a "natural" to a functionally "unnatural" display-control relationship. Further studies of compensatory and following tracking and their transfer effects are needed to provide additional information.

## H. SUMMARY

Different groups of subjects performed successively on a following tracking task and a compensatory tracking task.

When compensatory tracking with the "unnatural" display-control relationship was practiced after previous practice on a following tracking, a positive transfer effect was observed. No transfer effects were noted in other combinations of the two tasks.

## REFERENCES

1. BROWN, G. E., JR., Generality of tracking training. OSRD Report 4508, NDRC Applied Psychology Panel Memo. No. 9, 1944.
2. CARTER, L. F., & MURRAY, N. L. A study of the most effective relationship between selected control and indicator movements. Chapter 10 in P. M. Fitts (*Ed.*), *Psychological research on equipment design.* Report No. 19, U. S. Army Air Forces Aviation Psychology Program Research Reports, Washington, D. C.: U. S. Government Printing Office, 1947.
3. EDWARDS, A. L. Experimental Design in Psychological Research. New York: Rinehart, 1950.
4. GREEN, R. F., ANDREAS, B. G., & SPRAGG, S. D. S. Interference effects in compensatory tracking (modified *SAM* Two-Hand Pursuit Test) as a function of the reversal of the display-control relationships. (In preparation.)
5. LEWIS, D., MCALLISTER, D. E., & ADAMS, J. Facilitation and interference in performance on the modified Mashburn apparatus: I. The effects of varying the amount of original learning. *J. Exp. Psychol.,* 1951, **41**, 247-260.
6. LINDQUIST, E. F. Goodness of fit of trend curves and significance of trend differences. *Psychometrika,* 1947, **12**, 65-78.
7. MCALLISTER, D. E. The effects of level of learning and of overlearning on proactive and retroactive facilitation and interference. ONR, Special Devices Center, Technical Report No. SDC-938-1-8, August 1950.

8. McAllister, D. E., & Lewis, D. Facilitation and interference in performance on the modified Mashburn apparatus: II. The effects of varying the amount of interpolated learning. *J. Exp. Psychol.*, 1951, **41**, 356-363.

9. Melton, A. W. (*Ed.*), Apparatus Tests. Report No. 4, U. S. Army Air Forces Aviation Psychology Program Research Reports, Washington, D. C.: U. S. Government Printing Office, 1947.

10. Mitchell, M. J. H., & Vince, M. A. The direction of movement of machine controls. *Quart. J. Exp. Psychol.*, 1951, **3**, 24-35.

11. Norris, E. B., Green, R. F., & Spragg, S. D. S. Compensatory tracking performance (modified *SAM* Two-Hand Pursuit Test) as a function of directions and planes of movement of the control cranks relative to movement of the target. (In preparation.)

12. Norris, E. B., & Spragg, S. D. S. Performance on a following tracking task (modified *SAM* Two-Hand Coordination Test) as a function of the planes of operation of the controls. *J. of Psychol.*, 1953, **35**, 107-117.

13. ———. Performance on a following tracking task (*SAM* Two-Hand Coordination Test) as a function of the relations between direction of rotation of controls and direction of movement of display. *J. of Psychol.*, 1953, **35**, 119-129.

14. Shephard, A. H., & Lewis, D. Prior learning as a factor in shaping performance curves. *Proc. Nat. Acad. Sci.*, 1951, **37**, 124-131.

15. Warrick, M. J. Direction of movement in the use of control knobs to position visual indicators. Chapter 9 in P. M. Fitts (*Ed.*), *Psychological Research on Equipment Design.* Report No. 19, U. S. Army Air Forces Aviation Psychology Program Research Reports, Washington, D. C.: U. S. Government Printing Office, 1947.

*Department of Psychology*
*University of Rochester*
*Rochester 3, New York*

Published as a separate and in *The Journal of Psychology,* 1954, 37, 185-197.

# TRANSFER EFFECTS IN FOLLOWING TRACKING (MODIFIED SAM TWO-HAND COORDINATION TEST) AS A FUNCTION OF REVERSAL OF THE DISPLAY-CONTROL RELATIONSHIPS ON ALTERNATE BLOCKS OF TRIALS*[1]

*Department of Psychology, University of Rochester*

B. G. Andreas, R. F. Green, and S. D. S. Spragg[2]

## A. Introduction

Experimental studies have been conducted in which particular aspects of a task have been changed, with alternate practice under two different conditions required (2, 3). In some of these studies the directional relationship of the movement of the controls to the resultant changes in the display has been altered from one condition to the other. (It should be noted that the experimental manipulation here is actually the variation of stimulus-response relationships although the human engineering term, display-control relationship, is used in this article.) The effect of a complete reversal of the direction-of-movement relationship between display and controls has been frequently investigated (1, 4, 5, 6, 7, 11, 12). In most of these investigations *S*s were required to make discrete motor responses rather than the continuous adjustive response that is typically found in tracking. Further, the alternation pattern was usually limited to reversal from a standard to a reverse task and back to the standard again. The effects of repeated reversals of display-control movement relations on performance of a continuous adjustment task has not been systematically studied by varying the length of the blocks of trials in the alternation pattern. The exception to this is the study by Spieth and Lewis (12) in which two tasks were practiced in unequal blocks of trials in order to determine what ratio of trials on the

---

*Received in the Editorial Office on November 10, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was conducted as part of a research program on human motor performance carried out under Contract N6onr-241, T.O. 6 between the University of Rochester and the Special Devices Center, Office of Naval Research, Port Washington, L. I., N. Y., to whom these results have been reported in Technical Report SDC 241-6-10.

[2] The authors wish to thank Dr. Alfred H. Shephard for helpful suggestions concerning the planning of the experiment. We also want to acknowledge the assistance of Mr. John Tomlinson who conducted the statistical analysis of the data.

reversed task to trials on the standard task would result in maximum retardation of learning of both tasks.

In the present study the modified *SAM* Two-Hand Coördination Test was used to investigate the effects on following tracking of changing from the "natural" to the "unnatural" display-control relationship for alternate blocks of practice trials, the length of the block of trials being varied from group to group but held equal for each task. A previous study (10) has shown that the best tracking performance on the modified Two-Hand Coordination Test occurs when the relationship of the controls to the display is the "natural" one, in which a clockwise rotation of the left crank moves the target follower to the right and a clockwise rotation of the right crank moves the follower away from the operator. Since these are the movements

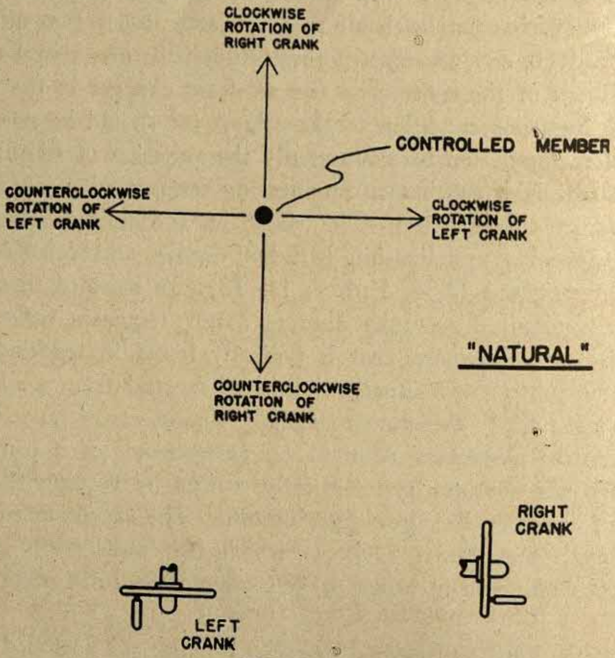

CLOCKWISE ROTATION OF RIGHT CRANK
CONTROLLED MEMBER
COUNTERCLOCKWISE ROTATION OF LEFT CRANK
CLOCKWISE ROTATION OF LEFT CRANK
"NATURAL"
COUNTERCLOCKWISE ROTATION OF RIGHT CRANK
RIGHT CRANK
LEFT CRANK

FIGURE 1

SCHEMATIC TOP VIEW OF APPARATUS SHOWING THE "NATURAL" DISPLAY-CONTROL RELATIONSHIP

The subject's position is at the bottom of the diagram. Controlled display member (black dot) is the target follower button. Each arrow indicates the resultant movement of this controlled member when a particular crank rotation is made. Intermediate directions of movement result when both cranks are turned simultaneously.

The "unnatural" condition involves exactly the opposite relations between direction of crank rotation and movement of the controlled display member.

of the controlled member which most persons expect when the cranks are thus rotated, and since these relationships correspond to those widely encountered in operating mechanical devices of all sorts, this arrangement has been designated as the "natural" one. Figure 1 illustrates the movements of the target follower that result from rotations of the two cranks under the "natural" condition. In the "unnatural" condition, the opposite rotation of each crank is required to obtain these movements of the follower. Under both conditions there was continuity between the plane of movement of each crank and the line of movement of the target follower that resulted from the rotation of the crank.

The data on transfer effects which are sought in this experiment have potential applicability to industrial and military situations in which a machine operator or a pilot, for example, may be called upon to operate alternately two devices which are essentially similar but which may be opposite in some display-control relationship. Of course, the periodicity of alternation would rarely be as regular in the real situation as in the experimental; nevertheless, the research findings on the interference or negative transfer effects under such conditions may suggest changes in training or operational procedures or in equipment design aimed at minimizing such interference.

## B. SUBJECTS

The subjects were 240 men taking basic training at a military installation. They were assigned to the eight different groups in the order of their appearance. None of the subjects had had previous experience with the apparatus.

## C. APPARATUS

The apparatus was a modified *SAM* Two-Hand Coördination Test (CM101B), described in detail in a previous study (9). By rotating the two control cranks the subject attempted to keep the target-follower button on top of a 5/8 inch diameter target which moved irregularly in the horizontal plane during each one-minute trial. The left crank rotated in a vertical plane parallel to the frontal body plane of the subject; the right crank rotated in a vertical plane perpendicular to this frontal body plane. A previous study (9) showed that time on target scores were higher with the cranks located in this combination of planes. Synchro linkages between cranks and the target follower button permitted *E* to reverse the display-control relationships as required.

## D. PROCEDURE

Eight groups of 30 *S*s were used, each *S* being given 20 one-minute trials on the modified Two-Hand Coördination Test. During the first 16 trials for the experimental groups, Groups I through VI, the "natural" and "unnatural" display-control relationships were alternated in blocks of one, two, four, or eight trials. For Groups I through IV, the initial block of trials was given under the "natural" condition, while the "unnatural" condition was experienced first by Groups V and VI. Groups VII and VIII, control groups, were given the "natural" or "unnatural" condition, respectively, for the first 16 trials. For purposes of intergroup comparison all groups were given four additional trials, in regular sequence, in which the "natural" condition was in effect on Trials 17 and 19 and the "unnatural" condition obtained on Trials 18 and 20. All subjects were informed whenever the experimenter reversed the display-control relationship. The pattern of alternation of conditions over the first 16 trials is indicated in Table 1.

*S*s were given general instructions for the task which were patterned after those given by Melton (8, p. 192). In addition, *E* stated: "After certain trials I am going to reverse the controls so that each crank must be turned in the opposite direction to get the same movement of the button that you were getting before the change. I'll tell you whenever I reverse the controls."

The two measures of performance taken on each trial were the total time on target measured with a Standard Electric Timer calibrated in milliminutes and the number of "hits" recorded on an electric counter.

TABLE 1

PATTERNS OF ALTERNATION OF DISPLAY-CONTROL RELATIONSHIP BETWEEN THE "NATURAL" (*N*) AND "UNNATURAL" (*U*) CONDITIONS IN THE SIX EXPERIMENTAL GROUPS AND TWO CONTROL GROUPS

| | Length of block | Pattern of alternation, first 16 trials* |
|---|---|---|
| *Experimental groups* | | |
| I | 1 | N-U-N-U-N-U-N-U-N-U-N-U-N-U-N-U |
| II | 2 | N-N-U-U-N-N-U-U-N-N-U-U-N-N-U-U |
| III | 4 | N-N-N-N-U-U-U-U-N-N-N-N-U-U-U-U |
| IV | 8 | N-N-N-N-N-N-N-N-U-U-U-U-U-U-U-U |
| V | 2 | U-U-N-N-U-U-N-N-U-U-N-N-U-U-N-N |
| VI | 8 | U-U-U-U-U-U-U-U-N-N-N-N-N-N-N-N |
| *Control groups* | | |
| VII | (16) | No alternation, *N* only |
| VIII | (16) | No alternation, *U* only |

*All groups received four additional trials, Trials 17 and 19 under the "natural" condition, Trials 18 and 20 under the "unnatural" condition.

## E. RESULTS

A primary interest in this study was the final level of performance attained during the experiment under the "natural" condition (Trials 17 and 19) and under the "unnatural" condition (Trials 18 and 20) after different patterns of alternation. Since these trials were actually the 9th and 10th with a particular condition or task, the performance of the control groups on Trials 9 and 10 was used as a standard in assessing the final achievement of the experimental groups.

Table 2 shows the final level of performance reached by the experimental groups under the "natural" and "unnatural" conditions. Also included are data from the control groups for the 9th and 10th trials under a particular condition.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF TIME ON TARGET SCORES IN MILLIMINUTES ATTAINED BY EACH GROUP ON THE NINTH AND TENTH TRIALS WITH A PARTICULAR DISPLAY-CONTROL RELATIONSHIP ($N = 30$ IN EACH GROUP)

| Group | "Natural" Mean | *SD* | "Unnatural" Mean | *SD* |
|---|---|---|---|---|
| I | 459.4 | 141.7 | 328.6 | 134.7 |
| II | 471.7 | 142.3 | 316.2 | 118.4 |
| III | 524.1 | 144.6 | 335.0 | 107.8 |
| IV | 514.7 | 160.8 | 358.3 | 138.3 |
| V | 526.7 | 132.8 | 367.3 | 107.6 |
| VI | 499.3 | 151.8 | 328.5 | 107.4 |
| VII | 518.6 | 153.4 | — | — |
| VIII | — | — | 428.9 | 165.4 |

Figure 2 shows the final performance level attained under the "natural" condition by the six experimental groups. Their mean scores did not differ significantly from that of the control group (Trials 9 and 10) which is represented by the horizontal dotted line in the figure. The data for the four groups that started with the "natural" condition were subjected to analysis of variance together with the control group data. For 4 and 145 degrees of freedom an *F* value of 1.15 was found, indicating that final achievement with the "natural" display-control relationship was not affected by the interspersed "unnatural" trials. In the case of the two groups which were given the "unnatural" condition first, inspection of the data served to indicate that their mean scores for final "natural" performance did not depart significantly from that of the control group.

Figure 3 shows the final level attained under the "unnatural" condition by the six experimental groups. Analysis of variance showed that the four groups which were initially given the "natural" condition did not differ from each

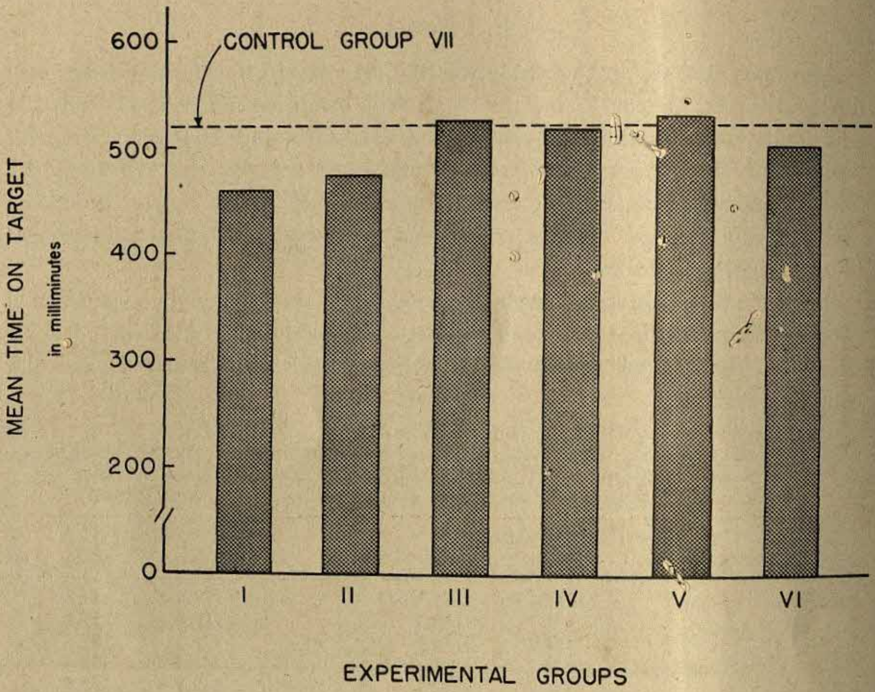


FIGURE 2

Mean Time on Target Scores Attained on the Ninth and Tenth Trials with the "Natural" Display-Control Relationship in Effect ($N = 30$ in Each Group)

other ($F$ was less than 1.0 for 3 and 116 degrees of freedom) so their data were combined. The grand mean for these four groups was then shown to be significantly lower than that of the control group (Trials 9 and 10) which is represented by the horizontal dotted line in the figure ($t = 2.88$, $df = 148$, $p < .01$). Separate comparisons of each experimental group mean with that of the control group gave similar results. This indicated that the final level of performance attained during the experiment with the "unnatural" display-control relationship was depressed by the preceding and interspersed "natural" trials.

The data for experimental Groups V and VI were also combined (since their means were not significantly different) and their grand mean was compared with that of the control group. Again a significant depression of the experimental mean was found ($t = 2.39$, $df = 88$, $p < .05$). When compared individually with the control group mean, the mean for Group

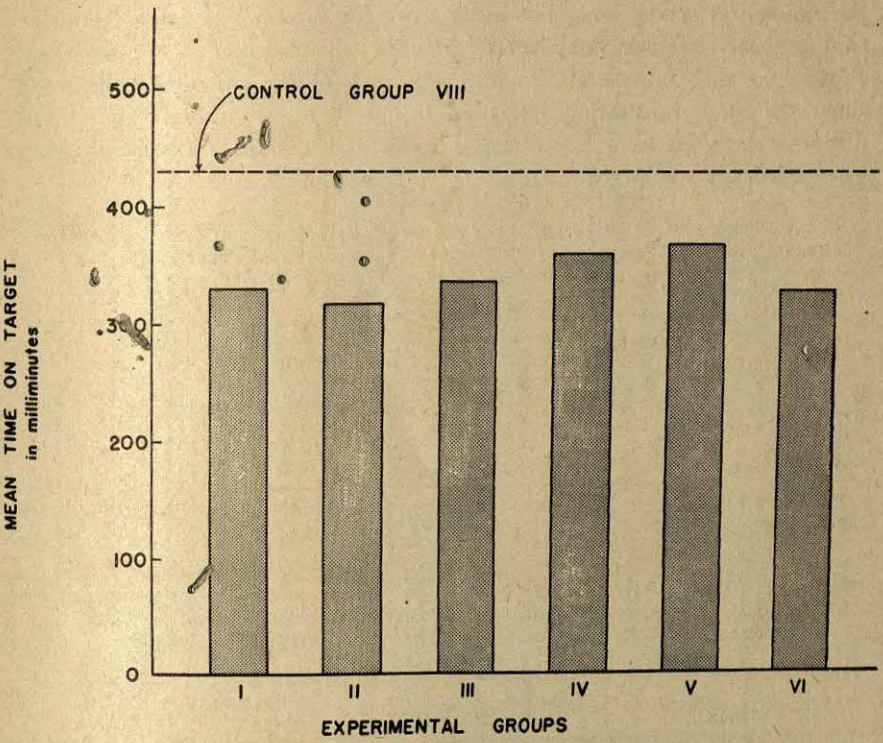


FIGURE 3
MEAN TIME ON TARGET SCORES ATTAINED ON THE NINTH AND TENTH TRIALS WITH THE "UNNATURAL" DISPLAY-CONTROL RELATIONSHIP IN EFFECT
($N = 30$ IN EACH GROUP)

VI was significantly lower; that for Group V was not quite significantly lower.

The depression of the final level attained by the experimental groups under the "unnatural" condition did not appear to be dependent upon (*a*) the size of the block of trials used in alternating the conditions nor (*b*) the condition which was experienced first in the alternation pattern since no significant differences were found among the means of the various groups differentiated by either of these factors.

A second major performance effect reflected in the data of this study was the transfer effect to the first trial with a particular display-control relationship following different amounts of practice with the opposite condition. Table 3 indicates the mean and *SD* of scores attained on the first "natural" trial by Groups V, VI, and VIII after 2, 8, and 16 trials, respectively, with the "unnatural" condition. The mean score for Group VII (first "natural" trial *without* previous "unnatural" practice) is presented for comparison. Practice on the "unnatural" condition in each of the amounts studied had about the same facilitating effect on initial "natural" performance. The difference between each experimental mean and that of the control group was statistically significant.

TABLE 3
MEANS AND STANDARD DEVIATIONS OF TIME ON TARGET SCORE IN MILLIMINUTES FOR INITIAL "NATURAL" TRIAL AFTER DIFFERENT AMOUNTS OF PREVIOUS PRACTICE WITH THE "UNNATURAL" DISPLAY-CONTROL RELATIONSHIP
($N = 30$ IN EACH GROUP)

| Group | Number of previous "unnatural" trials | Initial "natural" trial Mean | *SD* | *t* |
|---|---|---|---|---|
| V | 2 | 310.7 | 121.5 | *2.56 |
| VI | 8 | 303.3 | 139.0 | *2.21 |
| VIII | 16 | 302.2 | 132.4 | *2.24 |
| VII (control) | None | 220.8 | 144.7 | — |

*All *t* values indicate that each experimental group mean exceeded that of the control group at beyond the 5 per cent level of confidence.

TABLE 4
MEANS AND STANDARD DEVIATIONS OF TIME ON TARGET SCORES IN MILLIMINUTES FOR INITIAL "UNNATURAL" TRIAL AFTER DIFFERENT AMOUNTS OF PREVIOUS PRACTICE WITH THE "NATURAL" DISPLAY-CONTROL RELATIONSHIP
($N = 30$ IN EACH GROUP)

| Group | Number of previous "natural" trials | Initial "unnatural" trial Mean | *SD* |
|---|---|---|---|
| I | 1 | 164.5 | 116.3 |
| II | 2 | 199.1 | 117.6 |
| III | 4 | 242.9 | 104.6 |
| IV | 8 | 251.5 | 130.4 |
| VII | 17 | 152.8 | 83.9 |
| VIII (control) | None | 204.4 | 119.0 |

Table 4 indicates the mean and *SD* of scores attained on the first "unnatural" trial by Groups I, II, III, IV, and VII after 1, 2, 4, 8, and 17 trials, respectively, with the "natural" condition. The mean score for Group VIII (first "unnatural" trial without previous "natural" practice) is presented for comparison.

Figure 4 is a graphic representation of the means presented in Table 4. It shows that moderate amounts of prior practice under the "natural" condition tended to facilitate initial scores on "unnatural." Smaller or greater amounts of prior practice seemed actually to have an interference effect when conditions were reversed from "natural" to "unnatural." Analysis of variance established the fact that the five means represented in Figure 4 were signifi-

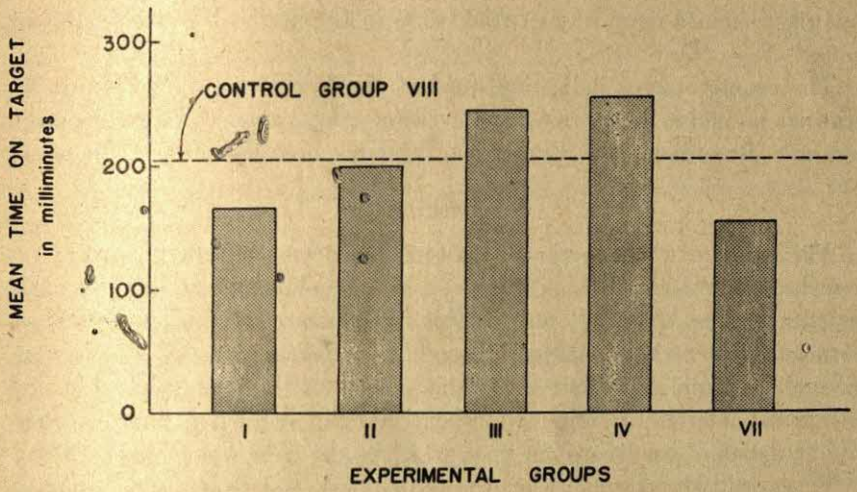


FIGURE 4
MEAN TIME ON TARGET SCORES ON INITIAL "UNNATURAL" TRIAL AFTER DIFFERENT AMOUNTS OF PRACTICE WITH THE "NATURAL" DISPLAY-CONTROL RELATIONSHIP ($N = 30$ IN EACH GROUP)

cantly different from each other ($F = 4.63$, $df = 4, 145$). Since these means were noted to be in the same rank order as the "natural" mean scores on Trial 1, it was considered possible that sampling differences among the groups had been responsible for the apparent differences. Consequently the data for the first "unnatural" trials were also subjected to analysis of covariance, using the Trial 1 scores as a control variable; this analysis yielded an $F$ ratio of 3.98 which for 4 and 144 degrees of freedom is significant at beyond the 1 per cent level of confidence. The result of this second analysis supported the conclusion that initial performance under the "unnatural" condition is a curvilinear function of the amount of prior "natural" practice. When compared by $t$-test with control data the mean score for Group VII was found to be significantly depressed ($t = 2.40$, $df = 118$, $p < .05$), indicating a negative transfer or interference effect. The mean for Group I showed a similar depression that did not quite reach significance ($t = 1.72$, $df = 118$, $p < .10$). The mean score for Group IV was found significantly higher than the control mean ($t = 2.02$, $df = 118$, $p < .05$), indicating a positive transfer or facilitation effect. For Group III the positive transfer effect fell slightly short of statistical significance ($t = 1.77$, $df = 118$, $p < .10$). Inspection of the data indicates that the net transfer effect in Group II approximated zero order. For all of these $t$-tests the control group mean utilized was the combined mean of first trial scores in Groups V, VI, and VIII with a total $N = 90$.

The counter scores indicating number of "hits" were found much less sensitive as indices of learning and transfer effects than the time on target measure. Accordingly the counter data have not been included in this report.

## F. DISCUSSION

The results of this experiment indicate that a transfer effect between two complex continuous adjustment tasks may be unidirectional. For example, practice on the "natural" task in this study depressed final tracking performance under the "unnatural" condition whereas practice on the "unnatural" task did not affect final achievement on the "natural." This unidirectional interference effect was apparently not dependent on the pattern of alternation of conditions nor upon which of the tasks was practiced first.

When initial performance on the second task was examined, a different pattern of transfer effects was observed. Initial performance under the "natural" condition was facilitated equally by three different amounts of prior practice on the "unnatural" task. The facilitation effect did not persist beyond three or four trials. Transfer from the "natural" to the first trial under the "unnatural" condition appeared as either a proactive facilitation or a proactive interference, depending upon the amount of the previous practice. It should be noted that the facilitation, where observed, was only temporary; as additional "natural" trials were interspersed the "unnatural" performance was depressed.

It was noted above that the final level of performance attained under the "unnatural" condition was not dependent on the *size* of the block of trials used in the alternation of conditions. Control over achievement in two such tasks was attained by Spieth and Lewis (12) who varied the *ratio* of trials on standard ("natural") and reversed ("unnatural") tasks with a turret tracking apparatus, finding a ratio of three reversed trials to one standard trial to be most effective in depressing or retarding performance on both tasks.

Results obtained by other experimenters when subjects were switched from the "natural" to the "unnatural" condition have not yielded the same curvilinear relationship of initial performance on the second task to practice on the first. For example, Lewis, McAllister, and Adams (4) found that different amounts of original learning on the standard Mashburn task ("natural") did not differentially affect the facilitation observed on the first trial of performance under the reversed ("unnatural") condition. Using a modified Two-Hand Coördination Test, with cranks located differently from those in the present experiment, McAllister (6) found that higher degrees of original learning on the standard task seemed to yield more proactive interference in the reversed task but this trend in the data was not found to be significant. The differences between the results of the earlier studies and those of the present experiment are probably due to differences in task and experimental procedures.

Even when considered by themselves, the results of the present experiment pose several problems of interpretation. The data are probably not extensive enough to warrant application of any formal theory of transfer, but a few assumptions, basically *ad hoc* in character, may be suggested for testing in subsequent studies. The first fact to be accounted for is the negative transfer effect which appears upon reversal to the "unnatural" condition, showing up initially in two groups and finally in all six experimental groups. This may be assumed to be due to a habit interference factor arising from practice under the "natural" condition. Upon reversal of the display-control relationships, *S* is required to make new (and opposite) crank-turning responses to "old" stimuli (the relative locations of target and target follower). Error-count data obtained in studies not yet published indicate that cranks are actually turned repeatedly in the wrong direction after reversal of conditions. These data support the assumption that it is response conflict or specific habit interference which gives rise to the negative transfer effects observed in the time-on-target scores for the "unnatural" condition.

Since moderate amounts of "natural" practice lead to a *positive* transfer effect later when the "unnatural" condition is initially encountered, it appears necessary to assume the operation of a second factor which is opposite to the specific negative factor. It seems reasonable here to postulate a positive factor, general task-familiarity, whose rate of acquisition is such that after moderate amounts of "natural" practice, it is sufficiently strong to counterbalance or even obscure the negative factor resulting from specific habit interference. This general factor may be considered as including, for example, familiarity with target course directions and velocities which is acquired and applied during practice with either the "natural" or the "unnatural" condition. To account for the results of this experiment it must be assumed that in one trial of "natural" practice a greater amount of the specific negative factor was acquired relative to the general positive factor so that a net negative transfer effect was observed when the task was changed to the "unnatural" condition. The two-trial acquisition of these opposed factors may be assumed to have yielded amounts of each which effectively cancelled so as to give rise to a net transfer effect of about zero magnitude. A further assumption about the growth of the hypothetical positive factor would logically be that it reaches an asymptote somewhere between eight and 17 trials so that the specific negative factor is in the ascendancy after 17 trials of "natural" practice and, it may be predicted, beyond that amount. The notion of deceleration of growth of the general positive factor, coupled with an assumed continuing growth of the habit interference factor, is also congruent with the failure of positive transfer effects to persist in this experiment as alternate practice continued under the opposed conditions.

To account for the fact that no significant negative transfer effects were observed in performance under the "natural" condition after various amounts of prior and interspersed practice with the "unnatural" display-control relationship, it may be assumed that the specific response habits under the "natural" condition are stronger than those acquired by the "unnatural" practice and thus more resistant to interference. Evidence for this assumption comes from the fact that scores under the "natural" condition are usually observed to be higher than those made on "unnatural" trials, and from the finding by Spieth and Lewis (12) that three trials of reversed practice had to be grouped with one trial of standard ("natural") practice in order to maximize negative transfer effects in the performance of each task. This assumption does not mean that performance under the "natural" condition was free from the negative habit factor but only that the level of this factor was always low enough to permit it to be masked or balanced by the general positive factor postulated above. It should be mentioned that the superiority of the "natural" condition is itself likely to be a transfer effect based on *S*s' previous experience with devices having these display-control relationships.

For practical purposes it is important to note that the "natural" condition—which yields the higher tracking scores—is also the one which was found in this study to have a final performance level which was not depressed by prior and interspersed practice on the "unnatural" condition. Thus "natural" display-control relationships can be recommended to the designer of equipment for their resistance to habit interference as well as for their initial facilitation of performance.

## G. SUMMARY

Six experimental groups were given different patterns of alternation between a "natural" and an "unnatural" condition of following tracking on a modified Two-Hand Coördination Test.

When the mean scores achieved on the 9th and 10th trials under the "natural" condition were compared with the appropriate mean from one

control group, no effect of alternating practice was observed. However, the comparable "unnatural" mean scores showed an interference effect when compared with their control group. The interference effect was apparently independent of the pattern of alternation.

Prior "unnatural" practice on 2, 8, or 16 trials gave the same amount of facilitation on the first "natural" trial. Prior "natural" practice for 1 or 17 trials tended to result in a negative transfer effect on the first "unnatural" trial, whereas the effect was positive facilitation after 4 or 8 "natural" trials. The facilitation effects in all instances were not persistent.

Some tentative assumptions regarding the acquisition of specific negative factors and a general positive factor are suggested to account for the results.

## References

1. Adams, J. A. The influence of the time interval after interpolated activity on psychomotor performance. *Res. Bull.* 52-11, Human Resources Research Center, Lackland Air Force Base, March, 1952.
2. Culler, A. J. Interference and adaptability. *Arch. of Psychol.,* 1912, **3**, No. 24.
3. Duncan, C. P., & Underwood, B. J. Transfer of training in motor learning as a function of degree of first-task learning and inter-task similarity. Technical Report 52-64, Wright Air Development Center, 1952.
4. Lewis, D., McAllister, D. E., & Adams, J. A. Facilitation and interference in performance on the modified Mashburn apparatus: I. The effects of varying the amount of original learning. *J. Exp. Psychol.,* 1951, **41**, 247-260.
5. Lewis, D., Shephard, A. H., & Adams, J. A. Evidences of associative interference in psychomotor performance. *Science,* 1949, **110**, 271-273.
6. McAllister, D. E. The effects of level of learning and of over-learning on proactive and retroactive facilitation and interference. ONR, Special Devices Center, Technical Report No. SDC 938-1-8, August, 1950.
7. McAllister, D. E., & Lewis, D. Facilitation and interference in performance on the modified Mashburn apparatus: II. The effects of varying the amount of interpolated learning. *J. Exp. Psychol.,* 1951, **41**, 356-363.
8. Melton, A. W. (*Ed.*) Apparatus Tests. Report No. 4, U. S. Army Air Forces Aviation Psychology Program Research Reports. Washington, D. C.: U. S. Government Printing Office, 1947.
9. Norris, E. B., & Spragg, S. D. S. Performance on a following tracking task (modified *SAM* Two-Hand Coördination Test) as a function of the planes of operations of the controls. *J. of Psychol.,* 1953, **35**, 107-117.
10. ———. Performance on a following tracking task (modified *SAM* Two-Hand Coördination Test) as a function of the relations between direction of rotation of controls and direction of movement of display. *J. of Psychol.,* 1953, **35**, 119-129.
11. Shephard, A. H. Losses of skill in performing the standard Mashburn task arising from different levels of learning on the reversed task. ONR, Special Devices Center, Technical Report No. SDC 938-1-9, August, 1950.
12. Spieth, W., & Lewis, D. The effects of alternating practice on the performance of two antagonistic motor tasks. ONR, Special Devices Center, Technical Report No. SDC 938-1-6, July, 1950.

*Department of Psychology*
*University of Rochester*
*Rochester 3, New York*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 199-223.

# ON THE USE OF AZAPORPHIN DERIVATIVES (PHTHALOCYANINES) IN STAINING NERVOUS TISSUE*[1]

*The Division of the Biological Sciences, University of Chicago*

HEINRICH KLÜVER AND ELIZABETH BARRERA

The white matter of the central nervous system has a characteristic fluorescence spectrum indicating the presence of porphyrins (19, 20). Chemical work has established the fact that the ether-soluble porphyrins occurring in the white matter of warm-blooded animals are protoporphyrin and coproporphyrin (19-21). One of us (22, 25) has suggested, on the basis of solubility characteristics, that it is the type III isomer of coproporphyrin which is present in the brain and spinal cord. Studies with a "fluorescence quenching" technique in Watson's laboratory have subsequently led to the same suggestion (4). More recently, Blanshard (3), working with fresh beef brain in the same laboratory, was able to obtain not only crystalline coproporphyrin III, but also a uro-type porphyrin the methyl ester of which melted at 273-4° C. The amount of the ether-insoluble uro-type porphyrin was even smaller than that of coproporphyrin III. Fluorimetric determinations have yielded 2 $\mu$g of coproporphyrin per 100 grams of beef brain (3), 3.56 $\mu$g/100 g. rabbit brain (4), and 2-3 $\mu$g/100 g. human brain (40). Our own results, when using the ethyl acetate-acetic acid method for extractions of mammalian brains, have in general corroborated these findings. Kôsaki (27) found values ranging from 1 to 7 $\mu$g per 100 g. brain tissue for the ether-soluble porphyrins in the cerebrum and cerebellum of avian and mammalian brains. These determinations mean that the total amounts of ether-soluble and ether-insoluble porphyrins do not exceed 10 $\mu$g/100 g. brain in normal animals if we consider only the metal-free porphyrins and ignore the porphyrins present in the hemoproteins of the brain. Although porphyrins occur in myelinated fiber masses, there is no evidence that they are constituents of particular tissue elements, such as the myelin sheaths. In fact, the view has been advanced that their occurrence in the white matter is correlated with the presence of oligodendroglia (21, 25). Pathological factors generally do

---

*Received in the Editorial Office on November 16, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research has been supported by a grant from the Commonwealth Fund. It has also been aided by the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

not produce striking increases in the porphyrin content of the brain and thus do not permit a greater degree of localization in the nervous system (25). It is worth mentioning that we found no higher porphyrin concentrations in brain tumors, e.g., in various gliomas, than in the central nervous system in general. In a recent summary of many years of work on factors influencing the concentrations of porphyrins and metalloporphyrins under normal and pathological conditions, Kôsaki (27) reached the conclusion that changes in the porphyrin content of organs are generally accompanied by changes in the concentrations of the iron-containing metalloporphyrins. He reported, for instance, that sulfonal derivatives produce striking changes in both the porphyrin and metalloporphyrin content of various bodily organs, tissues, and fluids; however, the highest concentration of brain porphyrin found by him was only 14 $\mu$g/100 g. cerebellum of guinea pigs (28). Further research will be necessary to determine whether even such an increase in ether-soluble porphyrins is the direct result of sulfonal poisoning. In general, the studies available at present leave little doubt that the amounts of porphyrins found in the brain under normal or pathological conditions are far too small to permit an exact localization of these compounds in the nervous system.

Since the low concentrations of naturally occurring brain porphyrins make it at present extremely difficult to determine their relations, if any, to structural elements, it seemed interesting to determine whether porphyrins introduced into the brain or added to brain slices or brain sections are taken up by particular cells or fibers. This paper deals chiefly with *in vitro* experiments, particularly with the effects of treating brain sections with porphyrins and related compounds. These procedures, although they did not disclose an affinity of porphyrins for particular fibers or cells, clearly demonstrated that porphyrins, under certain conditions, are chiefly deposited in regions of white matter. It was this observation which led us to develop a new method for the staining of myelin sheaths.[2]

Our preliminary tests showed that the fiber tracts of the brain, such as corpus callosum, anterior commissure, optic tract, external capsule, etc., are sharply outlined by red fluorescence if frozen sections of formalin-fixed monkey or rat brains are treated with alcoholic solutions of hematoporphyrin. More specifically, we passed the sections successively through distilled water and 50 per cent methanol, placed them in an acidified methanol solution of hematoporphyrin hydrochloride (containing 0.5 per cent glacial acetic acid), and then rinsed them in 50 per cent methanol before mounting from distilled water. This method yielded the same results when we used ethanol instead of methanol or employed hydrochloric acid instead of glacial acetic acid for acidification. The hematoporphyrin concentrations in these tests did not impart any visible red color to the sections, but produced a striking red fluorescence of all myelinated fiber tracts in marked contrast to the violet or weak reddish fluorescence of the gray matter. It deserves special mention that the fluorescence intensity of the granular layer of the cerebellar cortex was generally found to be somewhat greater than that of other cortical regions (Figure 1) and that the fluorescence spectrum of all fluorescent zones in sections thus treated did not represent an "acid" porphyrin spectrum, but one of Dhéré's Type I (13). (In this connection it is worth mentioning that the protoporphyrin-containing Harderian gland of the rat, when placed for several weeks in 10 per cent formalin, continues to exhibit a Dhéré Type I porphyrin fluorescence spectrum, with a principal emission band at 633 mμ, even if the formaldehyde solution shows an acid reaction.) Although hematoporphyrin, when used under the conditions just indicated, imparts marked fluorescence exclusively to myelinated fiber tracts, we obtained no evidence from fluorescence microscopy that this fluorescence was confined to particular structures in the white matter, such as myelin sheaths, glia cells, etc. Furthermore, we did not succeed in selectively staining such structures by means of metal complexes of hematoporphyrin. It is of some histochemical interest, however, that regions of the central nervous system which normally contain small amounts of free porphyrins are the ones that predominantly take up the fluorescent dye when free porphyrins are used for "staining." We obtained similar results in staining the porphyrin-containing root nodules of legumes with porphyrins or metalloporphyrins. As previously shown, the root nodules of leguminous plants contain not only hemoglobin, but also free porphyrins (22-24). The fluorescence spectrum of the infected hemoglobin-containing cells of such nodules generally exhibits an emission band at about 620 mμ. Several years ago, one of us (H.K.) succeeded in crystallizing the methyl ester of the legcoproporphyrin, a uro-type porphyrin, occurring in the root nodules of *Phaseolus vulgaris* (var. Red Kidney bean). If frozen sections of such legcoproporphyrin-containing root nodules are stained by our method for brain sections, the hematoporphyrin is found to stain the cells which normally contain legcoproporphyrin and leghemoglobin. For example, if relatively high concentrations of the zinc salt

---

[2] We are greatly indebted to Drs. Percival Bailey, Eleanor M. Humphreys, Stephen Polyak, Richard B. Richter, Gerhardt von Bonin, and Arthur Weil for supplying us with material for neurological staining and assisting us with helpful information in questions of histology and histopathology.

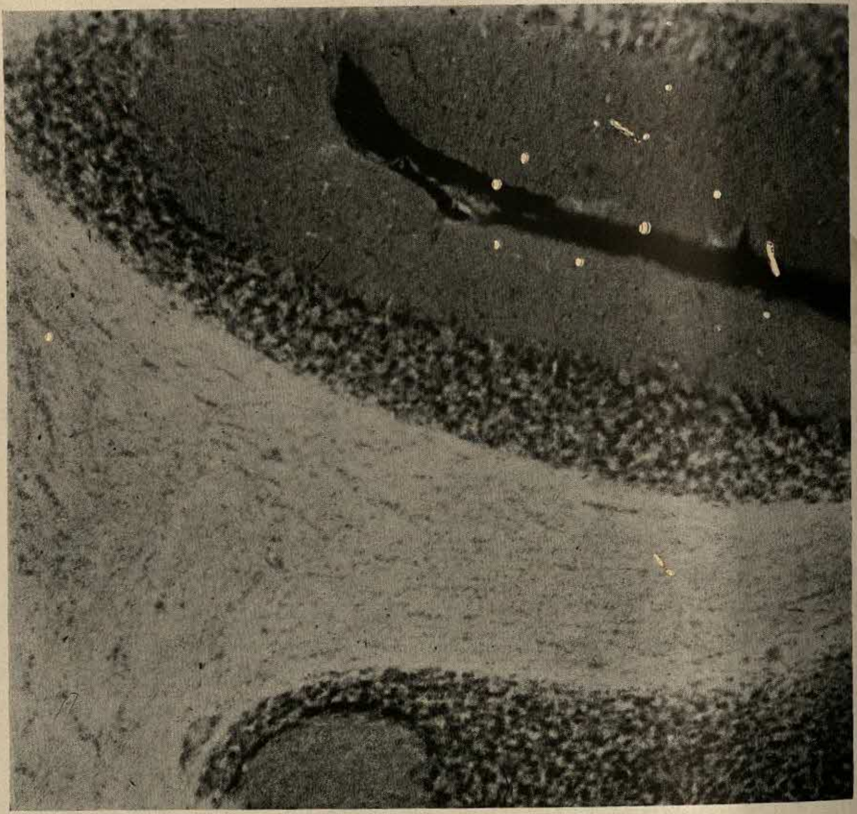


FIGURE 1
PHOTOMICROGRAPH OF THE FLUORESCENCE OF A 25μ FROZEN SECTION THROUGH CEREBELLUM OF THE RAT STAINED WITH AN ACIDIFIED ALCOHOLIC SOLUTION OF HEMATOPORPHYRIN HYDROCHLORIDE

of hematoporphyrin are used for staining, microspectroscopic examination of the nodule cells reveals the two-banded absorption and fluorescence spectrum characteristic of the zinc complex of hematoporphyrin.

To learn more about the possibilities of metal-free or metalloporphyrins for histological staining, we used acidified alcoholic solutions of hematoporphyrin to stain frozen sections of practically all organs. We shall briefly indicate the results obtained with organs of the rat. Fluorescence microscopy showed that the stain, judging by the intensity of porphyrin fluorescence in various tissues, was taken up by all muscular tissues, especially striated muscle, and, in varying amounts, by epithelial and glandular tissues. In general, connective tissue appeared optically empty. This method of staining seems particularly promising in the study of such endocrines as ovaries, testes, and adrenals.

In examining the adrenal glands of various mammals, including man, we generally found the porphyrinophilic zone of the adrenal cortex to be coextensive with the sudanophilic zone. This was also the case in adrenals removed from schizophrenic or carcinoma patients.[3] Examination of *unstained* frozen sections of human adrenals under the fluorescence microscope showed that capsule and cortex always comprised five distinct zones produced by variations in fluorescence colors. The porphyrinophilic zone was generally the same as the zone exhibiting yellowish-greenish fluorescence in unstained sections. (In this work we employed a mercury vapor lamp in conjunction with Corning filter No. 5874 as a source of illumination and made no attempt to exclude radiation from the red portion of the visible spectrum.)

While hematoporphyrin or naturally occurring free porphyrins have apparently heretofore not been used for purposes of staining, the use of another porphin derivative as a stain, namely, chlorophyll, dates back at least 40 years. In following the staining procedures suggested by previous workers (32, 33), we used chlorophyll to stain the adrenals, brain, and various other tissues. As Arndt (2) has already pointed out, however, chlorophyll has no advantages over Sudan and other fat stains except that it may be readily combined with other stains and that its green color may produce sharp contrasts not provided by the Sudan III and Sudan IV colors. Thus, a chlorophyll stain for fat may be successfully combined with Best's carmine stain for glycogen (33). It is worth mentioning that the cells of the root nodules of legumes may be as readily stained with chlorophyll as with porphyrin.

In view of the staining properties of porphyrins for the nervous system, the question may be raised whether there exist tetrapyrrolic ring compounds related to porphyrins which would be of greater value in neurological staining. Free porphyrins, while exhibiting an affinity for the white matter of brain and spinal cord under certain conditions, can so far not be successfully used for selectively staining tissue constituents, such as myelin sheaths, glia cells, or fibers. It must be recalled at this point that different porphyrins are distinguished only by various side chains in the $\beta$-positions of the pyrrole rings in porphin. Compounds exhibiting properties differing in many respects from those of the porphyrins are obtained by substitutions in, what

---

[3] We wish to thank Dr. Charles B. Huggins for kindly supplying us with portions of 18 human adrenals for histological and chemical work.

Fischer and collaborators (16) have called, the meso-positions of the porphin ring. For instance, when replacing one, two, or four of the CH groups, linking the pyrrolic rings, by tertiary nitrogen, various azaporphyrins (previously called imidoporphyrins) are obtained, such as mono-, di-, and tetra-azaporphyrins. Such a replacement of the methene bridges of the porphin skeleton by nitrogen atoms has marked effects on solubilities, HCl number, and the character of the absorption and fluorescence spectra (16). For example, as compared to the porphyrins, the distribution of intensities in the absorption spectrum is fundamentally changed even in the monoazaporphyrins, i.e., absorption band I is much stronger and band IV much weaker although the number of bands remains the same. Some azaporphyrins show a green color and no fluorescence in concentrated hydrochloric acid (15); their copper complex salts cannot be split even by concentrated sulfuric acid. The various azaporphyrins that have been synthesized, although not readily available, form an important link between the porphyrins and various other compounds with four benzopyrrole nuclei in which the pyrrole rings are condensed with benzene rings in their $\beta$-positions and in which the four methene groups joining the pyrrole nuclei are replaced by tertiary nitrogen. Fortunately, many of these latter compounds, i.e., the tetrabenzotetrazaporphins which have been named the phthalocyanines, are commercially available. They are of great interest to all investigators concerned with the chemical and physical aspects of porphyrins and hemoproteins (14, 17, 29, 30, 35, 38, 43, 44). As Wyman (43) has pointed out, it is somewhat paradoxical, in view of the complexity of the molecules involved, that the X-ray analysis of the structure of free and metal phthalocyanines "represents the first instance of the direct establishment of the structure of an organic compound from X-ray data." In fact, the chief evidence for the resonance structure of porphyrins has come from the study of phthalocyanines (29). Metal phthalocyanines representing every group in the periodic table have been synthesized (18). Many of the metallic derivatives have been found to function as catalysts in certain oxidation reactions, iron phthalocyanine being generally outstanding in this respect (7-10).

The highly colored blue to green phthalocyanines are said to represent "a truly startling innovation in the field of synthetic organic pigments" (31) and their discovery has been hailed as "by far the greatest contribution to the field of color" since Perkin's discovery of mauve about a century ago (34). Investigators concerned with industrial applications have emphasized that these pigments are characterized not only by excellent fastness to light and a truly remarkable resistance to heat and the action of most chemicals, but also by great beauty and intensity of the shades (11, 12, 18, 31, 34, 39). Among the metallic derivatives only the copper derivative, which must be considered the foremost blue pigment now available, has attained commercial importance. Phthalocyanines are being used for coloring lacquers, paints, plastics, cloth, rubber, leather, rayon, printing inks, etc. They give stable solutions in concentrated sulfuric acid and other strong acids, but are insoluble in water and nearly all organic solvents.

In view of the solubilities of the phthalocyanines, we tested the suitability of their sulfonated products in histological technique. We discovered that one of the commercially available products, namely, *Luxol fast blue MBS* (du Pont), which is an alcohol-soluble amine salt of sulfonated copper phthalocyanine, gives excellent results in staining myelin sheaths and that, in this respect, it is superior to Pontamine fast turquoise 8GL conc. 150 per cent (du Pont), another commercially available sulfonated copper phthalocyanine. In further experimenting with Luxol fast blue MBS, we found (*a*) that good myelin sheath preparations may be obtained without preliminary mordanting and (*b*) that a Nissl stain and a myelin stain may be combined in the same section. The details of the method for the combined staining of cells and fibers in the nervous system have recently been described (26). The only staining solutions used are an acidified alcoholic solution of Luxol fast blue MBS and an acidified aqueous solution of cresyl violet. The method requires only formalin fixation and is applicable to frozen and paraffin sections as well as to loose or affixed celloidin sections. It stains myelin sheaths and neurons as well as the nuclei of the three types of glia cells and differentiates the different layers of the medium-sized and larger blood vessels along with the blood cells. In brain sections, it generally furnishes beautiful cell-fiber preparations in which the reddish color of the Nissl-stained cells (particularly with the use of Grübler's *Cresylechtviolett*) provides a striking contrast with the greenish-blue color of the myelinated fibers. The photomicrographs in Figures 2-8A serve to illustrate some of the results of staining the central and peripheral nervous system of different animals. In Figure 8A the structure of the myelin sheath (as seen in fibers of the non-glial segment of the oculomotor nerve) represents merely one of various structures that can be seen when staining peripheral nerves by our method. In fact, all the structures figured by Spielmeyer as "normal structures of the myelin sheath" (Figure 61 in 36) have at one time or another been seen when applying our method to sections of peripheral nerves. Excellent cell-fiber preparations may be obtained even when old formalin-fixed material is used. We recently obtained such preparations when staining sec-

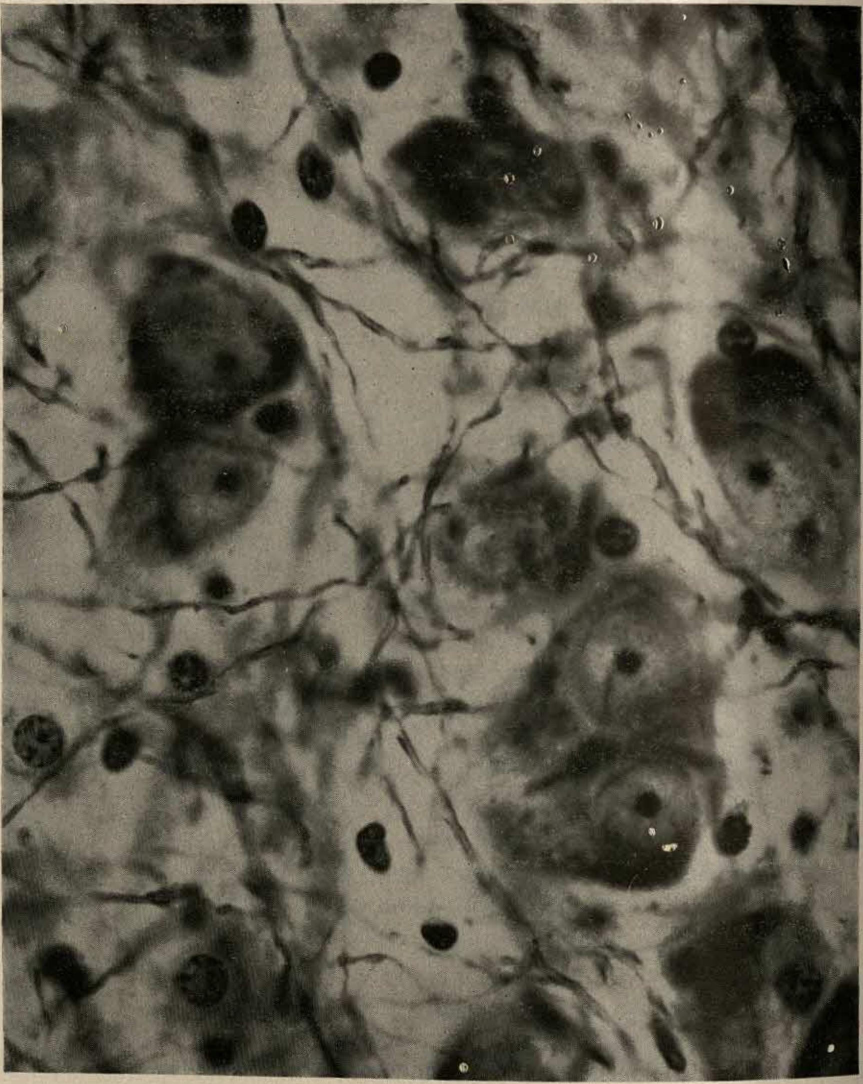


FIGURE 2
MYELIN SHEATHS, NEURONS, AND GLIA NUCLEI IN BRAIN STEM OF JAVA MONKEY: PHOTOMICROGRAPH FROM A 25μ FROZEN SECTION THROUGH THE MESENCEPHALON AT THE LEVEL OF THE INFERIOR COLLICULUS
(KLÜVER-BARRERA STAIN, x1000)

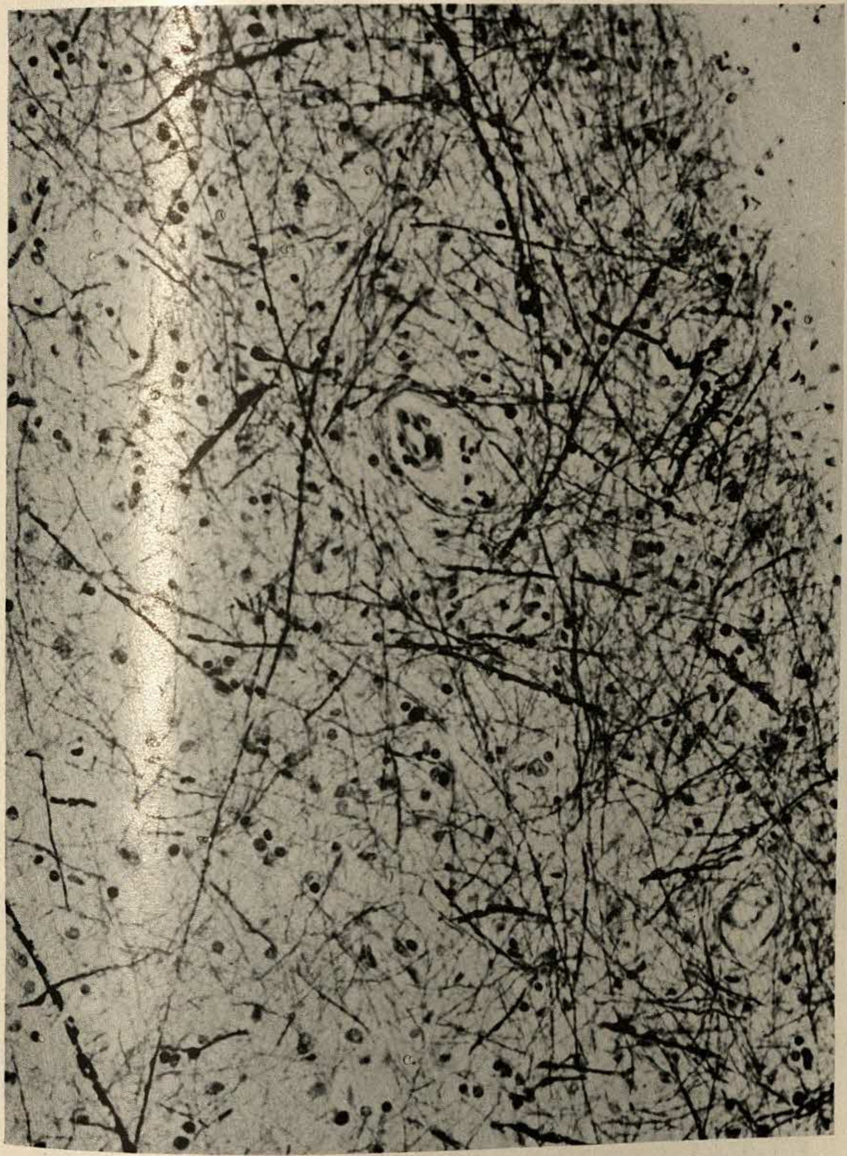


FIGURE 3
MOLECULAR LAYER OF CEREBRAL CORTEX IN MAN: PHOTOMICROGRAPH FROM A 15μ PARAFFIN SECTION (KLÜVER-BARRERA STAIN, x250)

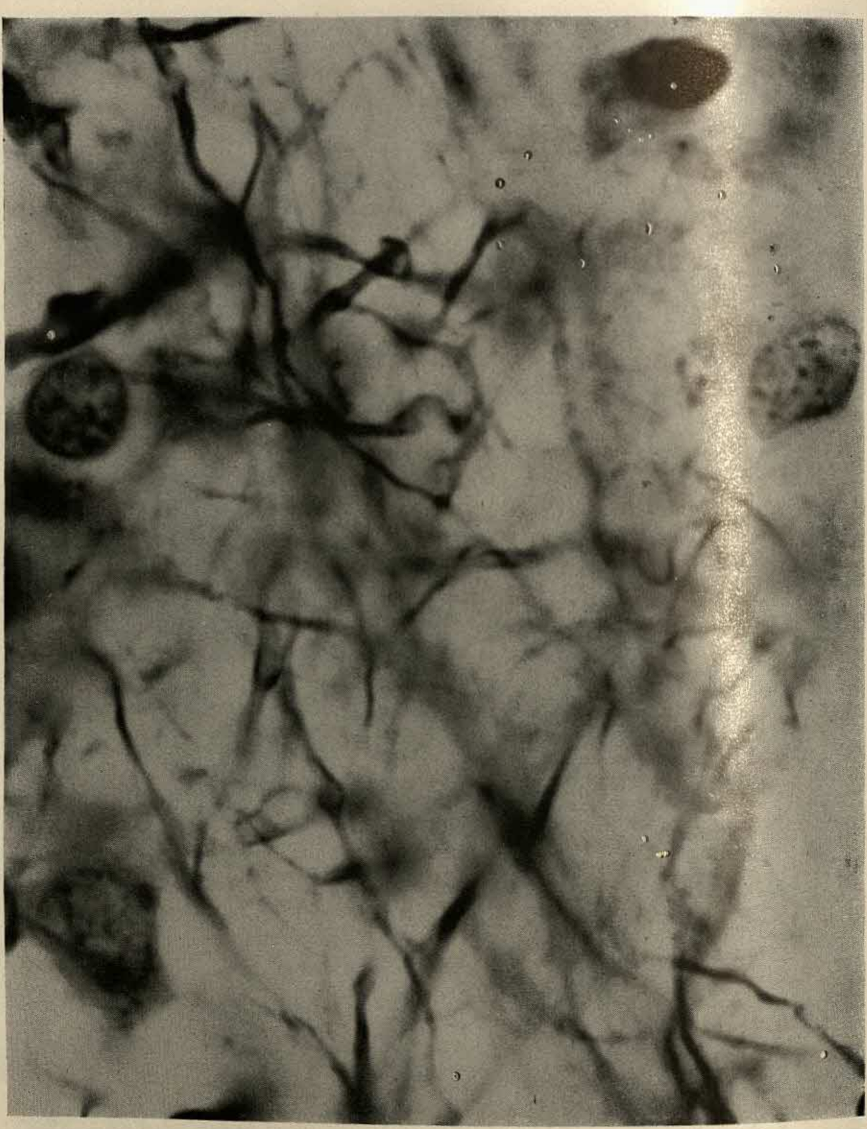


FIGURE 4
HIGHER MAGNIFICATION OF CELLS AND FIBERS FROM FIGURE 3, x2200

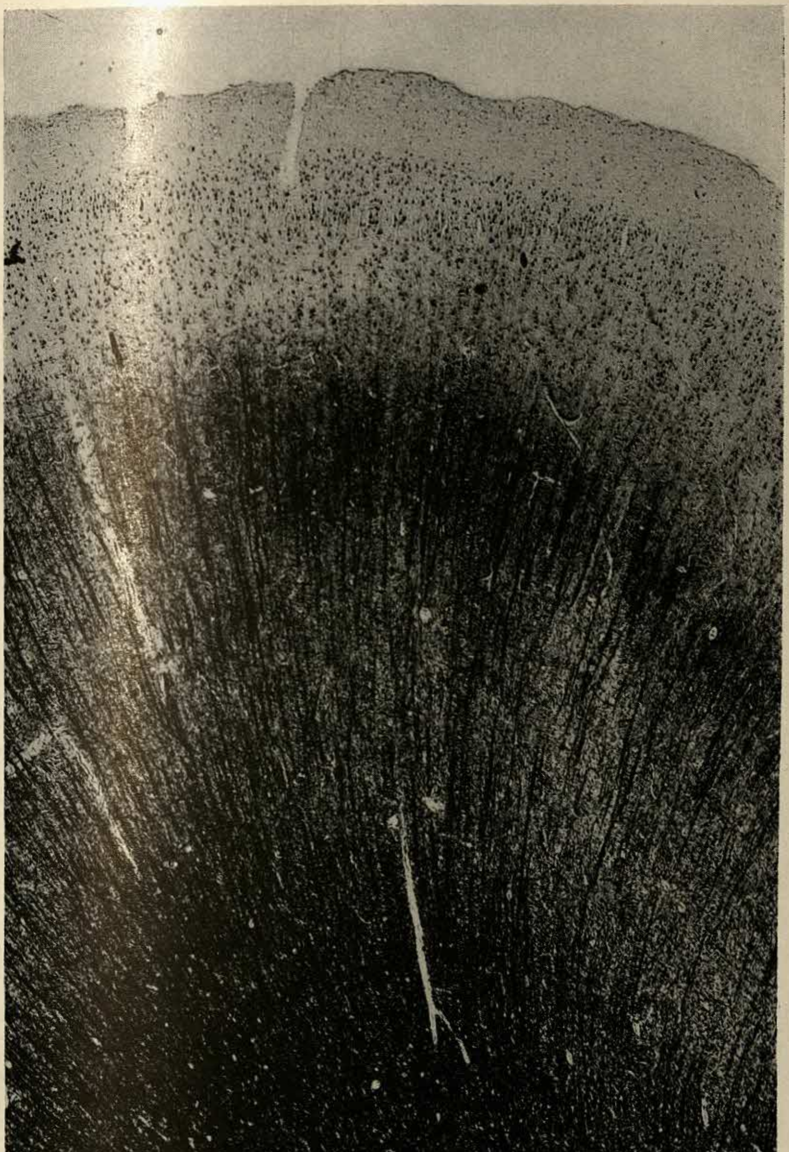


FIGURE 5
FRONTAL CORTEX OF GORILLA ("BUSHMAN"): PHOTOMICROGRAPH FROM A 15μ PARAFFIN SECTION (KLÜVER-BARRERA STAIN, x32)

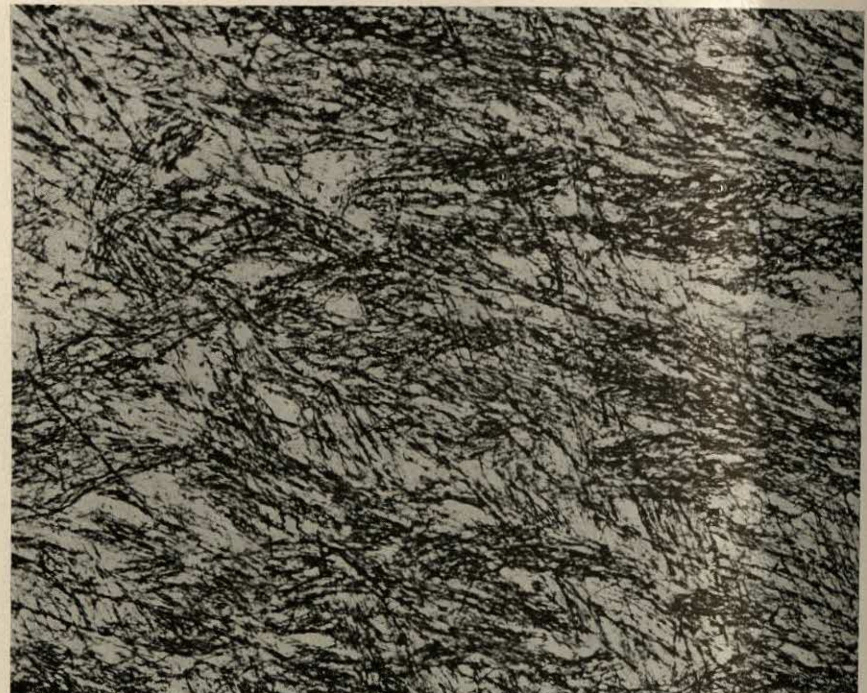


FIGURE 6
PHOTOMICROGRAPH FROM A 15μ PARAFFIN SECTION THROUGH THE MEDULLARY SUBSTANCE OF THE FRONTAL LOBE OF GORILLA ("BUSHMAN") (KLÜVER-BARRERA STAIN, x120)

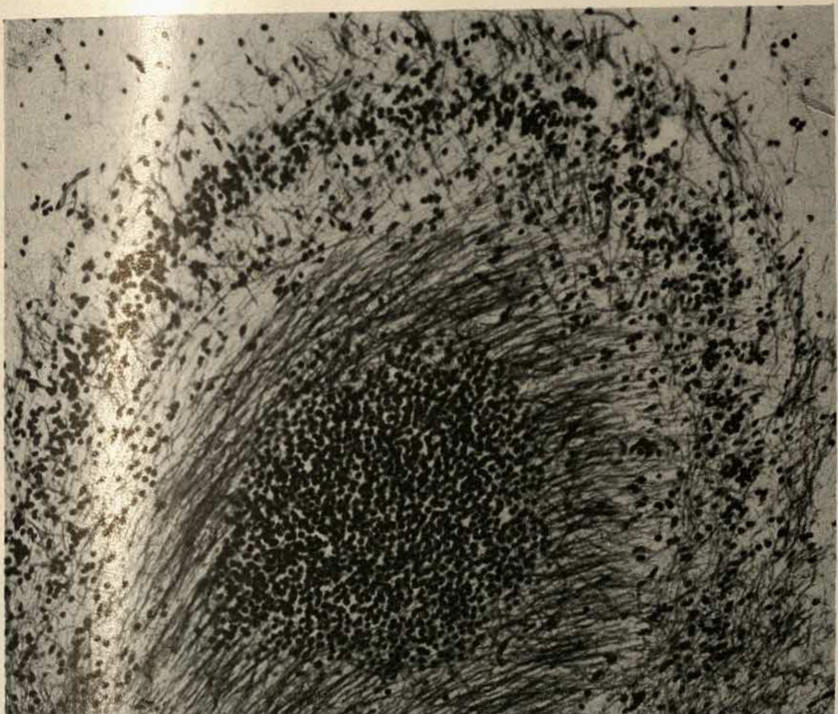
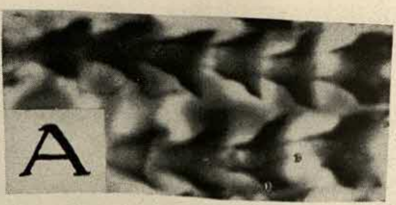


FIGURE 7
Photomicrograph from a 25μ Frozen Section Through Brain of the Leopard Frog
(Klüver-Barrera Stain, x145)



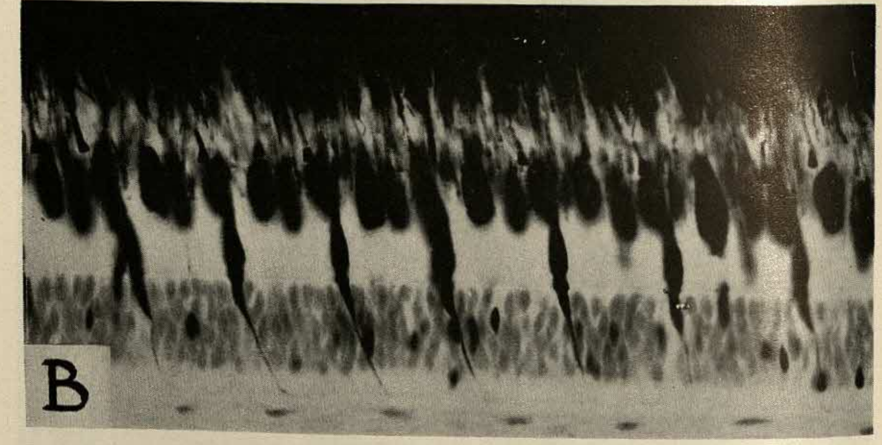


FIGURE 8A
Myelin Sheaths of Fibers of the Oculomotor Nerve in Man: Photomicrograph from a 16μ Celloidin Section (Klüver-Barrera Stain, x1000)

FIGURE 8B
Photomicrograph from a 14μ Celloidin Section Through the Teleost Retina (Brown Trout) (Klüver-Barrera Stain, x450)

tions of the brain of a gorilla ("Lady Congo")[4] that had been in formalin for about five years.

Luxol fast blue MBS may be used not only in neurological staining, but also in staining extraneural organs and tissues. We found that sections of various organs in crustaceans, amphibians, birds, and mammals may be successfully stained by simply employing our method (26), or modifications of our method, for cell-fiber preparations of nervous tissues. In some instances we used fixing fluids other than formalin. As a counterstain, we found acid fuchsin to be particularly valuable in making preparations of extraneural tissues. On the basis of work done so far, Luxol fast blue MBS appears especially promising in staining procedures for the retina, cochlea, pituitary, testes, ovaries, and adrenals (cf. Figures 8B and 9).

Further research on possible uses of sulfonated copper phthalocyanine in histological techniques may be facilitated by the availability of sulfonated products. They are marketed not only by American, but also by British and German concerns (in England, for instance, by Imperial Chemical Industries Ltd.). Copper phthalocyanine is generally considered the best commercially available pigment for the three-color process in color printing. Its structure is shown in Figure 10. It may be regarded as a chelate structure; there is retention of copper even when the compound is dissolved in concentrated sulfuric acid. None of the commercially available phthalocyanines can be considered chemically pure. The properties of Luxol fast blue MBS, therefore, are important. Of special interest in this connection are its absorption characteristics. The absorption curves in Figures 11 and 12 are based on measurements with a Beckman photoelectric quartz spectrophotometer. The staining solution which we used for nervous tissues was made by dissolving one gram of Luxol fast blue MBS in 1000 cc. of 95 per cent alcohol and adding 5 cc. of 10 per cent acetic acid. The absorption curve of a 0.001 per cent acidified alcoholic solution of Luxol fast blue MBS is shown in Figure 11 as compared to that of a 0.001 per cent solution of the compound in 95 per cent alcohol in Figure 12.

For visual measurements of absorption spectra, we employed a Gaertner L230 precision wave-length spectrometer. A 0.1 per cent Luxol fast blue MBS solution of 1 cm. thickness in 95 per cent alcohol showed a broad absorption band at 725-547 m$\mu$. A 0.01 per cent solution showed absorption bands at: I. 690-645 m$\mu$, II. 606-596 m$\mu$; order of intensities: I, II. A 0.001 per cent solution exhibited absorption maxima at: I. 666.3 m$\mu$, II. 600 m$\mu$.

---

[4] We are grateful to Dr. W. J. S. Krieg, Northwestern University, for putting tissue blocks from this brain at our disposal.

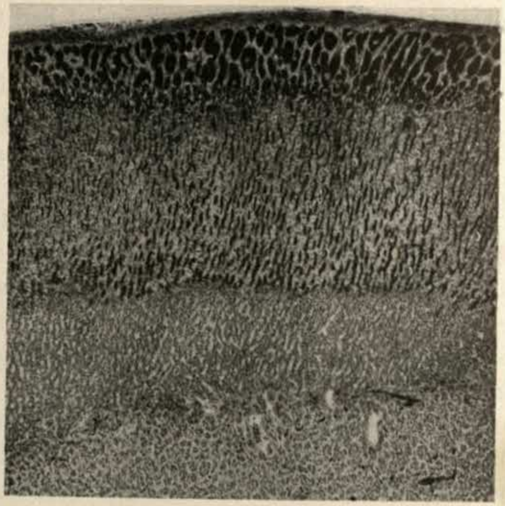


FIGURE 9
Photomicrograph from a 12μ Celloidin Section of the Adrenal Gland of Rhesus Monkey (Klüver-Barrera Stain)

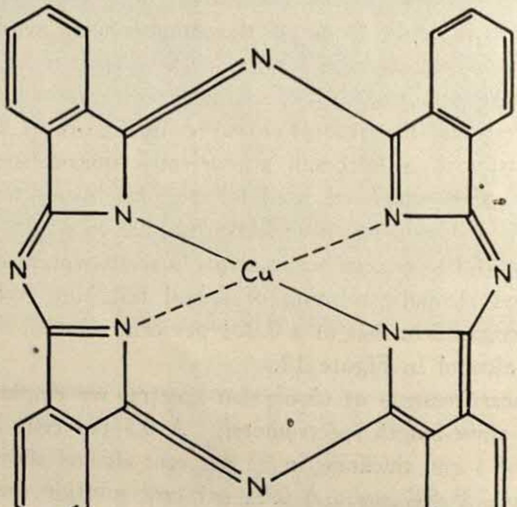


FIGURE 10
Copper Phthalocyanine

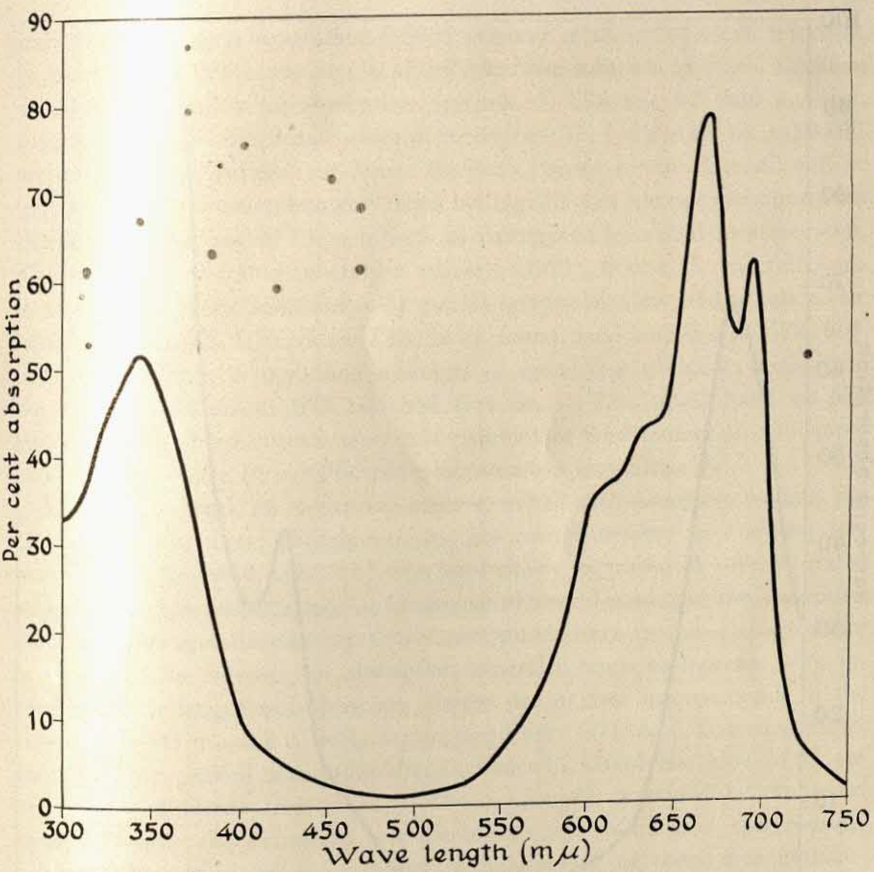


FIGURE 11
ABSORPTION CURVE OF LUXOL FAST BLUE MBS IN ACID 95 PER CENT ALCOHOL: CONCENTRATION, 0.001 PER CENT

Although a 0.0001 per cent solution showed no color, absorption band I, at about 666 mμ, could be seen. When we left a 0.1 per cent Luxol fast blue MBS solution standing for seven months at room temperature and in darkness and then made up solutions of different concentrations for measuring absorption spectra, we found that the passage of time had produced practically no changes in the absorption characteristics of the solutions (0.1 per cent solution: 724-547 mμ; 0.01 per cent solution: I. 690-648 mμ, II. 605-597 mμ; 0.001 per cent solution: I. 667 mμ, II. 601.2 mμ).

The addition of acetic acid to an alcoholic solution of Luxol fast blue MBS produced definite changes in the absorption spectrum. The acidified

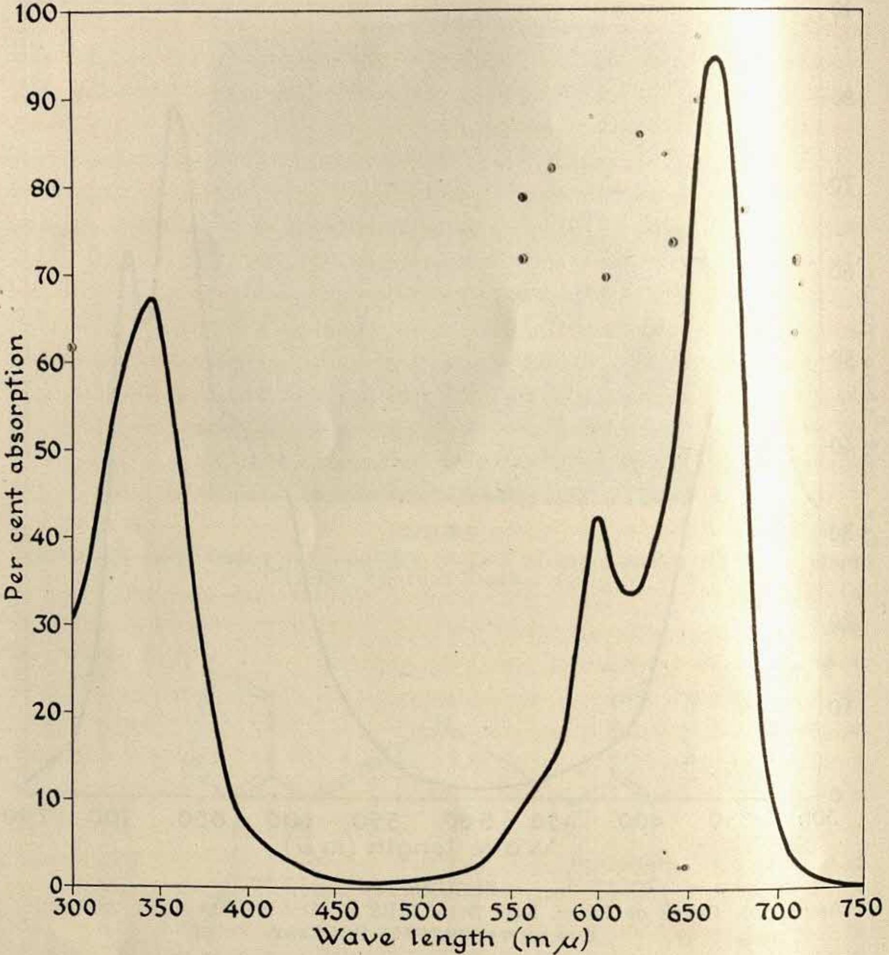


FIGURE 12
Absorption Curve of Luxol Fast Blue MBS in 95 Per Cent Alcohol: Concentration, 0.001 Per Cent

alcoholic staining solution previously referred to showed a broad absorption band at 740-547 mμ. An acidified 0.01 per cent solution exhibited three absorption bands: I. 704-647 mμ, II. 635-625 mμ, III. 606-596 mμ; order of intensities: I, III, II. An acidified 0.001 per cent solution exhibited absorption maxima at: I. 696 mμ, II. 670.5 mμ; order of intensities: II, I. Absorption band II was asymmetrical. When we estimated the position of the center of the band instead of the position of maximum absorption, we found the center of band II located at about 665 mμ. Absorption bands I and II could be seen in acidified 0.0001 per cent solutions of 1 cm. thickness. In examining a 0.001 per cent acidified alcoholic solution of 5 cm. thickness we found the following absorption maxima: I. 696 mμ, II. 666 mμ, III. about 632 mμ, IV. 601.3 mμ; order of intensities: II, I, IV, III. In a 0.00001 per cent acidified solution of 5 cm. thickness, bands I and II could still be seen with a hand spectroscope. Brief boiling did not change the absorption characteristics of any of the solutions as ascertained by visual measurements. The fact that we found absorption maxima at 696, 666, 632, and 601.3 mμ in acidified alcoholic solutions is of special interest in view of the results obtained by Stern and Pruckner (38) who found band maxima at 698, 664, 633, and 602 mμ in quinoline solutions of metal-free phthalocyanine, with the strongest bands at 698 mμ and 664 mμ. The small band at 675 mμ recorded by these authors was, in the opinion of Anderson *et al.* (1), probably due to a trace of metallic phthalocyanine as impurity.

Most of the work on absorption spectra, not of sulfonated compounds, but of metal-free or metal phthalocyanines, has been hampered by the very low solubilities. Anderson *et al.* (1) who used chiefly saturated or supersaturated solutions of chloronaphthalene or bromonaphthalene found that the absorption spectra of the metallic and metal-free compounds have certain characteristics in common: the intensity of absorption generally tends to increase with increasing wave-length and there are always one or two intense bands in the region 700-600 mμ and a fairly strong band near 600 mμ. It seems, therefore, that the general absorption characteristics of Luxol fast blue MBS are not greatly different from those of the compounds examined by Anderson *et al.* Despite the extremely low solubilities, we found that spectroscopically the absorption bands of copper phthalocyanine or metal-free phthalocyanine are as sharp and well-defined, for instance, in α-chloronaphthalene, as those of porphyrins in organic solvents.

Published data indicate that at low temperatures the bands in the absorption and fluorescence spectra of porphyrins shift towards shorter wave-lengths and become even sharper than at room temperature. At the same time there occurs a splitting of some of the absorption or emission bands into two or more bands (6, 22). The spectra of most colored substances, including those of chlorophyll derivatives, do not exhibit narrow bands even at low temperatures (6). The results we obtained in measuring the absorption spectra of Luxol fast blue MBS solutions at —195° C. suggest that lowering of the temperature produces effects similar to those observed in porphyrin spectra. For example, a 0.001 per cent solution in absolute alcohol exhibited at —195° C. the following absorption maxima: Ia. 671 mμ, Ib. 657 mμ, II.

about 632.6 mμ, III. 597.5 mμ. The minimum between bands Ia and Ib was located at about 663.2 mμ. At room temperature the absorption spectra of Luxol fast blue MBS in 95 per cent alcohol and in absolute alcohol were practically indistinguishable. The absorption spectrum of an acidified 0.001 per cent Luxol fast blue MBS solution in 95 per cent alcohol showed at —195° C. maxima at: Ia. 669 mμ, Ib. 659 mμ, II. about 630.8 mμ, III. about 597.8 mμ. The minimum between bands Ia and Ib was located at 664 mμ. The lowering of the temperature led to an almost complete disappearance of the 696 mμ absorption band.

The absorption spectra of greenish-blue microscopic sections of various organs stained according to our method (26) generally exhibited two bands in case no counterstain was employed. For example, the regions of white matter in paraffin, celloidin, or frozen sections of the human brain exhibited absorption maxima at 678 mμ and 597 mμ. Frozen sections of monkey brains showed bands at: I. 696-671 mμ, II. 622-572 mμ. In all sections of nervous and other tissues, band II was much broader than band I. In view of the absorption characteristics of the Luxol fast blue MBS stain, bluish illumination or daylight should preferably not be used for microscopy. It deserves special mention that, under certain conditions of illumination, sections stained with Luxol fast blue MBS show a striking color change. For instance, 4 μ paraffin sections of the brain may appear red instead of greenish-blue when examined under dark field illumination. Such a change in color is accompanied by a change in spectral characteristics.

In experimenting with Luxol fast blue MBS as a possible stain for myelin sheaths we discovered that acid alcohol as a solvent for the dye was preferable to either 95 per cent or absolute alcohol (26). Furthermore, in long-continued attempts to bring out cellular elements or improve the myelin stain, we used Luxol fast blue MBS in combination with numerous other stains, such as cresyl violet, thionin, toluidine blue, methylene blue, anilin blue, neutral red, methyl green, phloxine, acriflavine, phosphine, eosin, acid fuchsin, basic fuchsin, azure II, orange G, amaranth, hematoxylin, carmine, and orcein. These tests led to the conclusion that the cresyl violet—Luxol fast blue MBS combination is superior to any other combination in furnishing a Nissl stain and an excellent myelin stain in the same section. Only when using cresyl violet as a counterstain did we find, for instance, that Nissl granules, paranucleolar bodies (5), pigments in nerve-cells, and glia nuclei stained as well as, or even better than, in ordinary Nissl-stained preparations. No good cell-fiber preparations, however, could be obtained by using the cresyl violet of the Allied Chemical & Dye Corporation (National Aniline

Division) as a counterstain; by far the best results were obtained with Grübler's *Cresylechtviolett.* The Coleman and Bell domestic product was fortunately found to be a satisfactory substitute for the German product. It should be stressed that the counterstaining with cresyl violet, aside from bringing out the cellular elements, produces not only more sharply defined myelin sheaths, but may also modify the color of the myelin sheaths. At present we do not know why in some cell-fiber preparations the sheaths are purple rather than greenish-blue or why, in others, only the sheaths in certain fiber tracts or portions of certain fiber tracts assume a color different from that of the Luxol fast blue MBS stain.

Further research will also be necessary to determine why, in neurological staining, Luxol fast blue MBS may color not only the myelin sheaths, but also, under certain conditions, some of the cellular elements. In some of the sections stained according to our method, certain neurons and glia nuclei may exhibit the greenish-blue color of the dye. At times, only the nucleoli or the nuclei of certain nerve-cells show this color. No doubt, the affinity of certain cells or cell constituents for the stain is chemically interesting, but the histologist must be aware of one of its consequences, namely, that the regions with phthalocyanophilic cellular elements will show deviations from the regular Nissl picture if counterstained with cresyl violet. Although we have not succeeded in specifying the conditions responsible for the phthalocyanophilic character of certain cellular elements, we have made a number of observations bearing on this problem. A comparison of frozen, celloidin, and paraffin sections generally shows that an affinity of certain cells for the Luxol fast blue MBS stain is more likely to be found in frozen and particularly in celloidin sections than in paraffin sections. In fact, our method for paraffin sections should be used whenever it is desired to produce cell-fiber preparations with no or only slight deviations from the regular Nissl picture. A study of celloidin or frozen sections not counterstained with cresyl violet indicates that only cells in certain cortical layers, groups of neurons in certain nuclei, and sympathetic ganglion cells tend to show the characteristic greenish-blue color of the dye. The phthalocyanophilic elements in isocortical and allocortical regions are in most instances represented by pyramidal cells or at least by cells with a roughly triangular outline of the perikaryon. Furthermore, there is generally an increased affinity of certain nerve-cells and glia nuclei for the dye if, following the primary formalin fixation, the staining with Luxol fast blue MBS is preceded by treatment of the sections with 3 per cent potassium bichromate. Chromate treatment may thus alter the stratification of the cerebral cortex. For instance, in cell-fiber preparations of the striate cortex of macaques, a conspicuous and sharply outlined sublayer in layer V may appear which is not seen in ordinary Nissl preparations and which is produced by the marked phthalocyanophilia of certain cells in the fifth layer.

It also remains for future studies to determine the factors favoring the staining of the fibers of astrocytes in certain regions of the central nervous system. Although chromation following the primary formalin fixation appears to be one of the factors favoring the appearance of such fibers, sections not submitted to chromate treatment may also show the sharply defined, frequently green-colored fibers of astrocytes. The color of these fibers, unlike the color of certain myelin sheaths, is rarely modified by counterstaining with cresyl violet.

The staining of myelin sheaths by Luxol fast blue MBS raises the question as to the particular constituents of the sheaths colored by the dye. To throw some light on this question, formalin-fixed frozen sections were placed at room temperature (generally for 24 hours) in acetone, chloroform, benzene, xylene, ether, pyridine, or 2.5 per cent HCl before staining them for myelin sheaths or for both myelin sheaths and cells. None of the solvents entirely abolished the staining of myelin sheaths by Luxol fast blue MBS; recognizable sheaths could be found in all sections. Even preliminary treatment of the sections with mixtures of various solvents, such as ethanol-$H_2SO_4$ or ethanol-ether-HCl, did not prevent the coloring of the sheaths by the dye. In all instances, however, the staining of the sheaths was greatly inferior to that generally achieved by our standard technique. On the other hand, placing the brain tissue in lipid solvents, prior to the formalin fixation, completely abolished the staining of the sheaths. For instance, no evidence for a staining of sheaths could be obtained when rabbit brains, prior to formalin fixation, were placed for 48 hours at room temperature, or at 60° C., in acetone or, successively, in acetone, alcohol, ether, and pyridine. The evidence was likewise negative when the tissues were treated with various mixtures, such as ethanol-$H_2SO_4$, ethanol-ether-$H_2SO_4$, ethanol-chloroform, or pyridine-hydrazine hydrate, before fixing them in formalin. As regards future work, the most important finding in these experiments was that Luxol fast blue MBS, after preliminary extractions of the tissues with acetone and/or alcohol, stains axis-cylinders, at least in certain regions of the central nervous system.

In writing the history of his famous myelin stain in 1896, Weigert commented on the many inquiries he had received as to "how one finds such methods" (41). He tells the story that the discovery of the tubercle bacillus, more specifically, the special staining methods used by Koch for demonstrating tubercle bacilli, started him on the road to his own method for staining myelin sheaths. We have previously pointed out that the discovery of naturally occurring porphyrins in the central nervous system led us to the use of *phthalocyanine* derivatives, that is, to compounds chemically related to the porphyrins. At the same time, however, another set of data forced us to consider the rôle of copper compounds in the staining of myelin. If slices of brain or muscle as well as certain other tissues are left in contact with copper, the surfaces in contact with the metal acquire an intense reddish fluorescence (19). Further research by one of us (H.K.) showed that intramuscular injections of aqueous solutions of copper acetate or copper sulfate produce a striking reddish-yellowish fluorescence in the muscles at the point of injection and, moreover, that merely placing certain tissues, including brain tissue, in physiological saline containing copper compounds may be sufficient for eliciting an intense reddish fluorescence. As regards nervous tissues, the interesting observation was made that brain and spinal cord differ from peripheral nerves in regard to the fluorescence colors acquired by them through contact with copper or immersion in solutions of copper compounds. Preliminary observations indicated that the reddish, orange, or yellowish hues present in the fluorescence colors of copper-treated brain slices are generally absent in the fluorescence colors of similarly treated peripheral nerves or non-glial segments of cranial nerves. It was this set of observations which induced us to utilize copper sulfate in staining procedures, first of all, for the myelin sheaths of peripheral nerves. Although the results obtained in these first attempts are of no special interest in this connection, they provided the incentive for making a further search for copper compounds to be used in staining myelin sheaths in the central nervous system. It is, therefore, perhaps not surprising that such a search finally led us to the sulfonated product of a *copper* phthalocyanine. At present there is little doubt that not only the copper derivative, but also other metal derivatives or phthalocyanine itself may, under certain conditions, be successfully employed in myelin staining.

Finally, it needs emphasis that our method, in contrast to most of the traditional myelin methods, does not require chromation. Weigert, in attempting to characterize the reactions involved in his staining procedure, used the language of his cousin, Paul Ehrlich, in considering the mordant an "amboceptor" which combines, on the one hand, with the dye and, on the other hand, with particular constituents of the myelin sheath (42). There is probably general agreement that the physico-chemical principles advanced by Weigert in trying to explain the remarkable success of his techniques, are insufficient or wrong (37). At the same time it may seem regrettable and even somewhat puzzling that, despite the great advances in all branches of chemistry since the '80's when Weigert developed his myelin technique, "patient empiricism" and "never-ending trial and error methods," to quote Spielmeyer (37), should still play such a great rôle in developing a successful stain.

## References

1. Anderson, J. S., Bradbrook, E. F., Cook, A. H., & Linstead, R. P. Phthalocyanines and associated compounds: Part XIII. Absorption spectra. *J. Chem. Soc.*, 1938, 1151-1156.
2. Arndt, H. J. Zum histologisch-färberischen Lipoidnachweis mit Chlorophyll. *Zsch. wiss. Mikrosk.*, 1924, **41**, 481-486.
3. Blanshard, T. P. Isolation from mammalian brain of coproporphyrin III and a uro-type porphyrin. *Proc. Soc. Exp. Biol. & Med.*, 1953, **82**, 512-513.
4. Chu, E. J.-H., & Watson, C. J. An experimental study of the cerebral coproporphyrin in rabbits. *Proc. Soc. Exp. Biol. & Med.*, 1947, **66**, 569-570.
5. Coidan, R. S. The paranucleolar bodies in spinal neurons of mammals. *J. Comp. Neurol.*, 1952, **97**, 61-66.
6. Conant, J. B., & Kamerling, S. E. Studies in the chlorophyll series: VII. Evidence as to structure from measurements of absorption spectra. *J. Amer. Chem. Soc.*, 1931, **53**, 3522-3529.
7. Cook, A. H. Catalytic properties of the phthalocyanines: Part I. Catalase properties. *J. Chem. Soc.*, 1938, 1761-1768.
8. ———. Catalytic properties of the phthalocyanines: Part. II. Oxidase properties. *J. Chem. Soc.*, 1938, 1768-1774.
9. ———. Catalytic properties of the phthalocyanines: Part III. *J. Chem. Soc.*, 1938, 1774-1780.
10. ———. Catalytic properties of the phthalocyanines: Part IV. Chemiluminescent reactions. *J. Chem. Soc.*, 1938, 1845-1847.
11. Cronshaw, C. J. T. Phthalocyanines. *Endeavour*, 1942, **1**, 79-83.
12. Dahlen, M. A. The phthalocyanines: A new class of synthetic pigments and dyes. *Industr. & Engineering Chem.*, 1939, **31**, 839-847.
13. Dhéré, C. La fluorescence en biochemie. Paris: Presses Univ. de France, 1937. Pp. XII+318.
14. Endermann, F. Betrachtungen über die Struktur der Imido-Porphyrine im Zusammenhang mit den Phthalocyaninen. *Zsch. f. physik. Chem.*, Abt. A, 1942, **190**, 129-173.
15. Fischer, H., & Endermann, F. Synthese des Tetra-imido-ätio-porphyrins. *Ann. Chem.*, 1937, **531**, 245-250.
16. Fischer, H., & Orth, H. Die Chemie des Pyrrols: II. Band. 1. Hälfte. Leipzig: Akad. Verlagsges., 1937. Pp. x+764.
17. Fischer, H., & Stern, A. Die Chemie des Pyrrols: II. Band. 2. Hälfte. Leipzig: Akad. Verlagsges., 1940. Pp. ix+478.
18. Haddock, N. H., & Linstead, R. P. Phthalocyanines. *Thorpe's Dictionary of Applied Chem.* (4th ed.), 1949, **9**, 617-620.
19. Klüver, H. Porphyrins, the nervous system, and behavior. *J. of Psychol.*, 1944, **17**, 209-228.
20. ———. Porphyrins and the central nervous system. *Fed. Proc.*, 1944, **3**, 25-26.
21. ———. On naturally occurring porphyrins in the central nervous system. *Science*, 1944, **99**, 482-484.

22. ———. On a possible use of the root nodules of leguminous plants for research in neurology and psychiatry (preliminary report on a free porphyrin-hemoglobin system). *J. of Psychol.,* 1948, **25**, 331-356.

23. ———. On naturally occurring porphyrins in the root nodules of leguminous plants. *Fed. Proc.,* 1948, **7**, 66.

24. ———. Isolation of legcoproporphyrin. *Fed. Proc.,* 1949, **8**, 86-87.

25. ———. Functional differences between the occipital and temporal lobes with special reference to the interrelations of behavior and extracerebral mechanisms. In *Cerebral Mechanisms in Behavior* (ed. by L. A. Jeffress). New York: Wiley, 1951. (Pp. 147-182.)

26. KLÜVER, H., & BARRERA, E. A method for the combined staining of cells and fibers in the nervous system. *J. Neuropathol. & Exp. Neurol.,* 1953, **12**, 400-403.

27. KÔSAKI, T. Studies on porphyrins and metalloporphyrins. *J. Mie Medical College,* 1951, **2**, 85-174.

28. ———. On the pathogenesis of porphyria, especially of the toxic porphyria after the administration of sulfonal derivatives. *J. Biochem.,* 1951, **38**, 317-323.

29. LEMBERG, R., & LEGGE, J. W. Hematin Compounds and Bile Pigments. New York: Interscience Publ., 1949. Pp. xxv+749.

30. LINSTEAD, R. P. Porphyrins, phthalocyanines, and allied macrocyclic pigments. *Ann. Rep. Prog. Chem.* (Chem. Soc. London), 1937, **34**, 369-389.

31. PRATT, L. S. The Chemistry and Physics of Organic Pigments. New York: Wiley, 1947. Pp. vii+359.

32. ROMEIS, B. Mikroskopische Technik. München: Leibniz Verlag, 1948. Pp. xi+695.

33. SCHMORL, G. Die pathologisch-histologischen Untersuchungsmethoden. Leipzig: Vogel, 1925. Pp. x+481.

34. SCULLY, S. A. The phthalocyanine pigments. *Interchem. Rev.,* 1944, **3**, 39-44.

35. SEXTON, W. A. Chemical Constitution and Biological Activity. New York: Nostrand, 1950. Pp. xxi+412.

36. SPIELMEYER, W. Histopathologie des Nervensystems: I. Band. Berlin: Springer, 1922. Pp. vi+493.

37. ———. Technik der mikroskopischen Untersuchung des Nervensystems. Berlin: Springer, 1930. Pp. vi+168.

38. STERN, A., & PRUCKNER, F. Zur Lichtabsorption der Imido-porphyrine. *Zsch. f. physik. Chem.,* Abt. A, 1937, **178**, 420-436.

39. VENKATARAMAN, K. The Chemistry of Synthetic Dyes. Vol. II. New York: Acad. Press, 1952. Pp. xv+705-1442.

40. WATSON, C. J., & LARSON, E. A. The urinary coproporphyrins in health and disease. *Physiol. Rev.,* 1947, **27**, 478-510.

41. WEIGERT, C. Technik. *Erg. d. Anat. u. Entwicklungsgesch.,* 1896, **6**, 3-25.

42. ———. Nervenfasern, Markscheide. In *Enzyklopädie der mikroskopischen Technik* (ed. by R. Krause. Vol. III, pp. 1622-1629). Berlin: Urban & Schwarzenberg, 1927.

43. WYMAN, J., JR. Heme proteins. *Adv. in Protein Chem.,* 1948, **6**, 407-531.

44. ZEILE, K., & SIEDEL, W. Pyrrolfarbstoffe. In *Physiologische Chemie* (ed. by B. Flaschenträger, Vol. I). Berlin: Springer, 1951. (Pp. 845-980.)

*Culver Hall*
*University of Chicago*
*Chicago 37, Illinois*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 225-231.

# ELECTROMAGNETIC AND ELECTROSTATIC FIELDS: A NEGLECTED AREA IN PHYSIOLOGICAL PSYCHOLOGY*

*Mendocino State Hospital, Talmage, California*

W. E. BINGHAM, JR.[1]

## A. THE PROBLEM

This article surveys some recent, little-known research on magnetic and electrostatic fields from two points of view: (*a*) their existence in the nerve impulse and (*b*) their effects when applied to the body. There is a need to evaluate these electrical fields because they are possibly useful to psychologists in theory and experiment and because they have commanded such meager interest compared, for example, to the flood of action potential studies of the nerve impulse.

At present our knowledge of how magnetic and electrostatic fields can be related to the nervous system lies largely at the physiological level, though incomplete in respects which may have psychological implications. Actually, psychologists have contributed to some of the basic research. This survey is not exhaustive but topical and devoted mainly to the period, 1940 to 1950. A discussion of both fields is included in one article only for the conveniences of their having received so little attention and of their being manifestations of an electrical state with parallel qualities.

According to the conventional definition, a *magnetic field* is the portion of space in which a magnet or a body carrying an electric current experiences a force (11). In an electro-magnetic field the current and magnetism are considered to be components of a unitary field. An *electrostatic field* refers to the space in which a charge (potential) would experience a force (11). Each field has qualities of intensity, direction, attracting and repelling effects, and facilitation or inhibition according to the medium that contains the field.

## B. ELECTROMAGNETIC FIELDS

Research in magnetism of potential interest to psychologists falls into at least two main areas: (*a*) the existence of magnetic fields in the body and

---

*Received in the Editorial Office on November 17, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author is grateful to Dr. L. E. Travis and Dr. N. D. Warren of the University of Southern California for their initial encouragement, and to Dr. J. A. Gengerelli of the University of California at Los Angeles for making available his work in this area.

(*b*) the effect of an externally applied magnetic field to the body. A corollary to the latter is the possibility of bodily sensitivity to magnetism.

### 1. *Existence of Magnetic Fields in the Nerve Impulse*

In spite of the mesmerists' failure to verify animal magnetism, it is now reasonable to explore for the presence of magnetic fields in the body originating from the nerve impulse in view of its nature as a bioelectric current. The principles of electromagnetism require that along with the presence of an electric current there exists also a magnetic field at right angles to and surrounding a net flow of current in one direction (1).

When considering this very problem, Gengerelli (6) predicted, however, that due to the form of the current, the external field would be almost completely cancelled by the internal magnetic field. "If, therefore, a suitable detecting instrument is placed outside the nerve itself, no magnetic field should be observed" (6, p. 190). A frog's sciatic was passed through the long axis of a toroidal coil. In keeping with the prediction, no current was induced in the detecting coil, within the limits of the instrument's sensitivity, as the impulse passed through it, and even when the coil and nerve were immersed in a saline solution. However, it is an unsolved problem whether or not the instrument was sufficiently sensitive.

This experiment by Gengerelli implies that an internal magnetic field may be considered. His test coil, though, was not designed to detect a field internal to the flow of depolarization current. Even more comprehensive studies like that of Lorente de No (15) have completely overlooked the magnetic field as a factor to be rigorously ruled in or out. Yet one of these studies (15, vol. 131) is now valuable for its estimate of impulse current, because current, not potential, establishes magnetic flux.

### 2. *The Effect of Externally Applied Magnetic Fields*

Three divergent lines of investigation will be reported. First, there is the phenomenon of phosphenes (visual sensations aroused by stimuli other than light). Barlow, Kohn, and Walsh (2) may be consulted for the bibliography on magnetic phosphenes. Beginning with D'Arsonval in 1896, physicists, physiologists, and psychologists have reported phosphenes and sometimes taste sensations when the head was immersed in a strong *AC* magnetic field (2). For example, Dunlap (5) lowered a coil around his head and noted that "the whole visual field quivered as if illuminated by a rapidly intermittent light" (5, p. 50). He interpreted this as an alternate intensification and inhibition of whatever was observed. He also noted the flicker did not affect positive and negative after-images in the central portion of the visual field. When a magnetic intensity of 5,400 ampere-turns was attained, he found that with his head inside the coil the effect was "intensely disagreeable." Dunlap felt that whether the effect was retinal or occipital remained for further research.

Thirty-six years later this study was made by Barlow, Kohn, and Walsh (2) when they concentrated up to 900 gauss through the temple and occipital regions by means of a pointed-core electromagnet. They found that a shimmering in the visual field, usually colorless, sometimes tinted blue or yellow, resulted only when the magnetism was focused through the temple and never through the occiput. Hence, after eliminating also the optic nerve as a factor, it was concluded that the origin of magnetic phosphenes is in the retina. Nevertheless, this study needs to be carefully criticized for the adequacy of the psychological controls because so many variables in perception were ignored.

In a different sort of inquiry a carefully controlled experiment by Gordon (9) did not substantiate Yeagley's (16) theory explaining the homing pigeon's ability to home by means of its sensitivity to geographic differences in the grid established by the earth's magnetic field and the Coreolis force from the earth's rotation.

There is a third avenue of interest. Although mesmerists were repudiated in their belief that magnetism could cure diseases, Hansen (10) recently tested the notion that a local oedema would arise after application of electromagnetism to an inflamed part. She compared two histamine puncture reactions below the right and left shoulder blades, one side subjected to a *DC* field intensity of 1,000 gauss and the other side (control) to no field. Using an *N* of 25 people, she claimed a relative inhibition of the histamine flush under the magnet. However, the injection procedure was not completely controlled and no reliability statistics were offered. For psychologists the significance of this experiment might lie, not in its medical aspect, but in possible clues to understanding the interaction of magnetism with the nervous system.

## C. ELECTROSTATIC FIELDS

### 1. *Existence of a Static Field from the Nerve Impulse*

Several studies have demonstrated an electrostatic field attending passage of the nerve impulse. Gengerelli (7) detected it in frog nerve, using an induction coil, and differentiated its discharge objectively from the depolarization current. Burr and Mauro (4) built on the demonstration by Lorente de No (15, vol. 132) of an electrostatic field in an infinite liquid volume conductor accompanying nerve impulses in the frog's sciatic. They verified an electrical field in the air surrounding the nerve trunk of a preparation.

### 2. *The Effect of Externally Applied Static Fields*

In the absence of known studies of neuropsychological effects of a static field upon the body as a whole without discharging to an external ground, it is known what happens when an electrostatic field is applied to a resting nerve. Gengerelli and Holter (8) placed a frog's sciatic nerve and gastrocnemious muscle between two condenser plates seven inches apart. With a potential of 10,000 volts between the plates at a frequency of 60 cycles stimulation varied in degree with the angle that the nerve made with the static lines of force, a parallel relationship increasing the reaction. Also, when the nerve was immersed in saline solution, the effect was eliminated. The latter effect suggests that such static fields may be unable to stimulate nerves *in situ*.

## D. SUGGESTIONS FOR RESEARCH

### 1. *Implication of Fields in the Nerve Impulse*

The probable absence of an external magnetic field in the nerve impulse raises the question about an internal field. If such a field exists, it may serve no more function than that of the magnetic field of an electric light filament. Observe, nevertheless, that concerning the rhythmic character of nervous activity, Brink (3) has reviewed rival theories which differ on whether or not the nerve membrane behaves as if it had an inductance factor. Now the actual presence of inductance would require the existence of magnetism.

Forasmuch as an impulse current spreads beyond the membrane, an internal magnetic field may have some unknown function in the cerebral cortex where fibers are highly concentrated and impulses may interact. The likelihood of psychological implications as an intervening variable may then be entertained. For example, would there be any consequences for the direct current theory of cerebral integration and perception proposed by Koehler (12) when the magnetic component is considered?

The dearth of theory and research about a magnetic aspect to the current of the impulse probably stems from an historical overemphasis on thinking in terms of impulse potential rather than current upon which magnetic flux depends.

The implications of static fields of impulses, such as by interaction between them, would seem to be minimized by the grounding effects of body fluids noted above (8).

## 2. *Implications of Externally Applied Fields*

What are the limits of visual (retinal) sensitivity to *AC* magnetic fields? Since autonomic effects may arise by subception (13), i.e., below the level of awareness, sensitivity to weaker fields might be studied by aid of the galvanic skin response. Why do strong magnetic fields of varying intensity arouse sensations from the retina? Is the answer, electromagnetic induction or rhythmic interaction with magnetic fields inside impulses which were already discharging from the retina?

Just what are all the effects of externally applied magnetism? Are the qualities of any sense perception changed quantitatively under conditions of first shielding out the earth's magnetic field and then steadily increasing the strength of the constant *DC* field in steps to extreme values? Under these conditions are cerebral and peripheral neural mechanisms affected in any fashion as shown by electroencephalogram and galvanic skin response patterns? How do the *AC* magnetic fields similarly applied affect these patterns?

At present popular professional opinion says that magnetism has no effect upon the body, but it must be borne in mind that the brain, as one part of the body, has compensatory mechanisms such as for keeping upright visual perceptions when the head is tipped to one side. Both subjectively and behaviorally such a mechanism, without refined observation, might obscure actual effects of magnetism upon cerebral processes.

Next, what are the consequences, if any, of magnetic fields upon impulses along a single fiber or tract? Our knowledge of the nerve impulse may be advanced by studies of the effect of applied magnetic fields of various strengths, directions, and forms upon the electrical properties of the single nerve impulse as well as the resting fiber. Applied magnetic fields might be used in searching for evidence of an internal magnetic component of the impulse, such as by observing the result upon an impulse as it passes into a strong, directed, and properly shaped field. Some studies might be aided by using the Lillie (14) iron wire model of the nerve. Since the permeability of the iron might complicate the test, a search might be made for non-magnetic wires.

If a magnetic field exists internal to the impulse, could strong, properly oriented, and externally applied magnetism stop a slow, weak impulse like that found in type *C* fibers? In this event, based no doubt upon the attracting or repelling quality of interaction between two magnetic fields, we might learn to develop an impulse inhibitor which could serve as an anesthetic or a block to muscular reaction, perhaps without affecting tissue.

Finally, some of the above questions could also be asked about static fields.

The sting of discharging static electricity is a common experience to many people after making friction with another surface. And discharging static potentials through a patient to a ground has long been used in electrotherapy to stimulate nerves, skin, and muscles. But the problem here is a different one. Is there any effect of static fields only as applied to the body as a whole or to the cerebrum or to a nerve impulse without discharging to an external ground?

Only by making these and similar explorations can we assess the value of magnetic and static fields for psychology. At present knowledge lies largely in the domains of physics and physiology. In fact, this suggests a lesson for the psychologist who would experiment in this area. The ideal professional equipment consisting of a physicist, a neurophysiologist, and a psychologist is highly recommended in order to give proper attention to the many-sided facets of a total problem.

## E. Summary

This survey of recent research on electromagnetic and electrostatic fields in relation to the nervous system reveals undeveloped areas of potential significance for psychology. The basic findings follow: (*a*) A magnetic field probably does not exist external to the current flow of the nerve impulse, but an internal field remains to be considered. (*b*) It has been verified that intense alternating magnetic fields when applied to the head arouse visual sensations, apparently based upon effects in the retina rather than the occipital lobe or elsewhere. (*c*) An electrostatic field accompanies passage of the nerve impulse. (*d*) Intense electrostatic fields applied to a nerve preparation may arouse impulses. Suggestions for research point imaginatively to further knowledge of the impulse, medical aids, sensory functions, and possible cerebral mechanisms.

## References

1. Attwood, S. S. Electric and Magnetic Fields. (3rd ed.). New York: Wiley, 1949.
2. Barlow, H. B., Kohn, H. I., & Walsh, E. G. Visual sensations aroused by magnetic fields. *Amer. J. Physiol.*, 1947, **148**, 372-375.
3. Brink, Jr., F. Excitation and conduction in the neuron. In S. S. Stevens (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951. (Pp. 50-93.)
4. Burr, H. S., & Mauro, A. Electrostatic fields of the sciatic nerve in the frog. *Yale J. Biol. & Med.*, 1949, **21**, 455-463.
5. Dunlap, K. Visual sensations from the alternating magnetic field. *Science*, 1911, **33**, 68-71.
6. Gengerelli, J. A. External field of the nerve impulse. *Proc. Soc. Exp. Biol.*, 1942, **51**, 189-190.
7. ———. Electrical field of the nerve impulse. *Proc. Soc. Exp. Biol.*, 1943, **52**, 189-190.

8. GENGERELLI, J. A., & HOLTER, N. J. Experiments on stimulation of nerves by alternating electrical fields. *Proc. Soc. Exp. Biol.*, 1941, **46**, 532-534.

9. GORDON, D. A. Sensitivity of the homing pigeon to the magnetic field of the earth. *Science*, 1948, **108**, 710-711.

10. HANSEN, K. M. Studies made for the end of finding objective expression of influence of magnetism on man and of ascertaining whether this influence is carried by way of the autonomic nervous system. *Acta Med. Scandinav.*, 1949, **135**, 448-457.

11. HUMPHREYS, R. F., & BERINGER, R. *First Principles of Atomic Physics.* New York: Harper, 1950.

12. KOEHLER, W. Relational determination in perception. In L. A. Jeffress (*Ed.*), *Cerebral Mechanisms in Behavior.* New York: Wiley, 1951. (Pp. 200-243.)

13. LAZARUS, R. S., & MCCLEARY, R. A. Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.*, 1951, **58**, 113-122.

14. LILLIE, R. S. The passive iron wire model of photoplasmic and nervous transmission and its physiological analogues. *Biol. Rev.*, 1936, **11**, 181.

15. LORENTE DE NO, R. A study of nerve physiology. *Stud. Rockefeller Inst. Med. Res.*, 1947, Nos. 131 and 132.

16. YEAGLEY, H. L. A preliminary investigation of a physical basis of bird navigation. *J. Appl. Phys.*, 1947, **18**, 1035-1063.

*425 West Church Street*
*Ukiah, California*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 233-241.

# A MODIFICATION OF THE THEMATIC APPERCEPTION TEST FOR NAVAL ENLISTED PERSONNEL (*N-TAT*)*[1]

*U. S. Naval Medical Research Laboratory, New London, Connecticut*

DENNIE L. BRIGGS[2]
*(Assisted by the Assessment Staff)*

## A. THE PROBLEM

Early in the fall of 1950, shortly after the outbreak of the Korean emergency, various members of the staff of the Neuropsychiatric Unit at the U. S. Naval Training Station, Newport, Rhode Island,[3] strongly realized the need for modified clinical methods for assessing various personality traits at deeper levels in a minimum amount of time. Demands of the Naval Service require that such techniques be in a form suitable for administration to large numbers of personnel by hospital corpsmen who have had little or no formal training in psychological procedures. Out of these earlier attempts at developing techniques for such purposes (1, 2) came the realization that valuable clinical insight could be gained from Thematic Apperception Test (*TAT*) protocols written by the recruits themselves from lantern slides of the original set of pictures by Henry A. Murray and from the modification by Thompson.

Further studies made at New London confirmed the feasibility of modifying this technique, both in terms of the test stimulus itself and in the mode of administration. The pictures were reduced in size to approximately three by four inches and printed on sheets of paper which could then be inserted into booklets (see Figure 1). Thus the *order of presentation* and the *length of the test* could be varied according to the current and changing demands of the military service. It was further felt that by modifying certain of the original pictures, and through the addition of others depicting situations involving Naval personnel and events peculiar to Navy life, Naval personnel would be able to identify more readily with the stimulus presented and thus more clearly reveal some of the more unconscious elements of their personalities which might have direct relevance to the military situation.

---

*Received in the Editorial Office on November 19, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Based on the Thematic Apperception Test by Henry A. Murray and the staff of the Harvard Psychological Clinic, and sponsored by the Bureau of Medicine and Surgery of the Department of the Navy, Project No. NM 003 041.13. The opinions or conclusions contained herein are those of the author and are not to be construed as official or reflecting the views of the Department of the Navy or of the Department of Defense.

Acknowledgment is especially made to Commander Gerald J. Duffner, MC, U. S. Navy, Officer-in-Charge of the Laboratory for his suggestions, especially with the three new plates presented here. The author would also like to express his gratitude to Henry A. Murray for his many suggestions and encouragement, to Christiana Morgan for her advice in redrawing the original *TAT* pictures, and to Ivan N. Mensh and Commander Robert R. Deen, MC, U. S. Navy, for their suggestions in presenting the modification in this paper.

[2] Now at the U. S. Naval Hospital, Oakland, California.

[3] Commander Robert R. Deen, MC, U. S. Navy, then senior psychiatrist; Lieutenant Commander Herman B. Molish, MSC, U. S. Naval Reserve, then senior clinical psychologist; and the author.

Several forms of the test have been experimented with, including the use of forms which contained only the original Murray set of pictures, sets using the Navy modification, and sets using combinations of all the above and some with pictures from the Thompson modification. The final form, consisting of 10 pictures which have been selected from over 60, is used extensively in the assessment program for the submarine forces (5), and is presented in this paper together with a breakdown of typical stories given by a group of enlisted submarine candidates to three of the new pictures drawn especially for Naval personnel.[4]

The present paper is thus limited merely to a presentation of the modifications and of the technique used. Other studies will present the modification for officer personnel; the scoring system developed;[5] and comparative studies using a representative sample of Naval personnel, psychotics, the severely maladjusted and disciplinary offenders. The present form of the test has been given to over 5,000 Naval enlisted personnel.

## B. METHOD OF ADMINISTRATION

The candidates are assembled in groups of any size, or are given the test individually. They are told that this is a study of the imagination. The instructor usually comments on the importance of the imagination in a person and its relevance to the submarine service. The candidates are assured that there are no right or wrong answers, and that the test is to help the doctors get better acquainted with them prior to their interview which usually follows the next day. The instructions on the test booklet are as follows:

---

[4] The original drawings for the three *TAT* pictures presented here was the work of Seaman Kenneth Pelto, U. S. Navy, assistant illustrator of the Naval Medical Research Laboratory.

[5] Currently being accomplished by Dr. Leonard Eron and Dr. Frank Auld, Yale University, under a collaborative contract.

> This is a test of your imagination. You will find a picture at the top of each page in this booklet. Your task will be to make up as dramatic a story as you can for each of these pictures. Describe what is happening in each picture. Tell what the characters are thinking and feeling, keeping in mind the following:
>
> 1. What has led up to this incident?
> 2. What is happening now?
> 3. What will be the outcome?

Candidates are then given approximately an hour to write their stories, or about five minutes per picture. Experience has shown that this is sufficient time for the average person. However, no pressure is exerted, and individuals may take as long as they desire to finish their stories.

The modification has also been used successfully for clinical patients who are not too disturbed to handle such procedures, and further clinical studies are in progress at a Naval Hospital to determine the test's feasibility with psychotic patients.[6]

## C. USE OF THE TEST

Of immediate practical importance to the Submarine Medical Officer, the *N-TAT* is used as an aid in the assessment interview. Other techniques used have been described elsewhere.[7] Prior to his use of the test, the medical officer has received a training course in submarine medicine, lasting approximately six months, which includes seminars and clinical demonstrations in psychiatric principles and in clinical psychology with emphasis on projective testing.[8] A considerable amount of time is spent on presenting case work-ups in an attempt to show the significance of the material elicited by the test. The medical officer is then given supervised practice in the total evaluation of submarine candidates.

The submarine medical officer reads each *N-TAT* booklet just before interviewing the candidate and jots down hunches he may want to explore further during the interview. No attempt is made to score or formally present the material in the stories. In cases where he is not able to obtain enough evidence to support or to refute his hunches, or when he suspects he may have a case presenting serious maladjustment, the medical officer may refer the candidate to the psychiatrist for psychiatric evaluation, or to the clinical psychologist for further studies.

---

[6] Being conducted by the author and with the coöperation of Dr. John T. Evans, Chief Clinical Psychologist, U. S. Naval Hospital, Chelsea, Massachusetts.

[7] Crissy, W. J. E., and associates. The interview: II. Aids to the interview—the confidential questionnaire; III. The submariner stereotype; IV. The reliability and validity of the assessment interview as a screening and selection technique in the submarine service. Naval Medical Research Laboratory Reports 211, 214, and 216, 1952-53.

[8] In addition to the staff of the Laboratory, outside civilian consultants present seminars to the medical officers, including one on the *TAT* by Dr. Henry A. Murray.

## D. The Present Test Series

The 10 plates now being used consist of modifications of seven of Murray's pictures, all of which have been redrawn for various reasons. Three new pictures have been added to this series. The presentation to follow will be concerned only with the three new pictures which have been especially drawn for the Navy series. Figure 1 presents the 10 pictures currently being used, in the order finally selected.

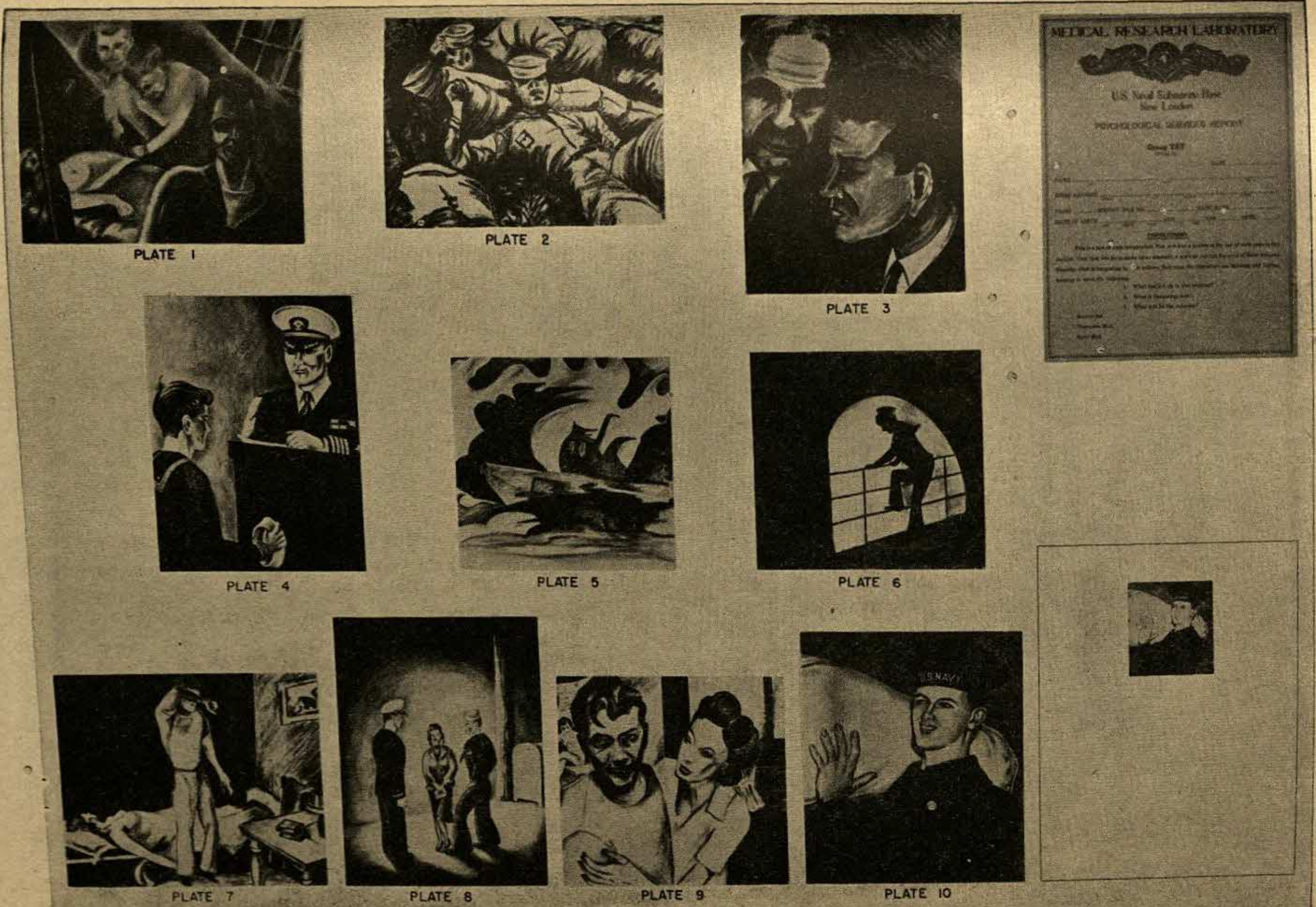
### 1. *Plate 4: The Sailor Before the Captain*

This picture was especially designed to elicit attitudes toward authority figures by Naval enlisted personnel and to determine how they perceive certain power relationships. We thus assumed that their present perceptions of authority might possibly be related to how authority was handled in their earlier years. From the responses to this picture, we are also able to gain more insight into their needs for power, for submission, or for punishment, by noting the harshness with which they deal with the figures in their stories.

The majority of the candidates interpret the picture as one where the sailor is at Captain's mast. Various methods are then used to deal with the sailor. Another smaller group interpret the picture as one where the sailor is at request mast to ask for some special consideration, and still another group see the situation as one where the sailor is being commended for some heroic deed. The frequencies given by a typical group of submarine candidates are given in Table 1.

### 2. *Plate 8: Sailor and Officer With Girl*

In this picture we were interested in finding out some of the attitudes of enlisted personnel toward officers, and of rôle-taking by the candidates. Here we also attempt to find how enlisted men perceive a social situation and their need for approval by their officers. Roughly, one-third of the candidates see this situation as the "eternal triangle," one in which there is active competition between the officer and the sailor for the attention of the girl. The manner in which the subjects settle the issue is of importance. Of those giving this theme, about 30 per cent let the officer win the girl, about 25 per cent let the enlisted man have the girl, and in a few number of stories,



NAVY THEMATIC APPERCEPTION TEST (N-TAT)

FIGURE 1

TABLE 1

PLATE 4: THE SAILOR BEFORE THE CAPTAIN ($N = 402$)

| | Frequency |
|---|---|
| *Theme 1: A Captain's Mast* | 100 |
| A. Sailor is guilty and receives specific punishment | |
| 1. Undo harshness implied in the Captain and/or the sentence; the sailor is made an example | 17 |
| 2. Is punished, but learns his lesson | 25 |
| 3. Is punished, but to no avail | 1 |
| | 143 |
| B. Sailor is guilty, but gets out of the situation | 6 |
| 1. Is given only a "warning" | 37 |
| 2. Is given a warning and learns his lesson | 20 |
| 3. Case dismissed; accused not guilty | 3 |
| 4. Although guilty, is commended | 1 |
| | 67 |
| C. Sailor is at Mast—outcome uncertain | 51 |
| 1. Presents more than one alternative outcome | 28 |
| 2. Outcome not mentioned, but learns his lesson | 16 |
| | 95 |
| *Theme 2: A Request Mast* | 4 |
| 1. Approved | |
| 2. Denied | |
| 3. Not certain | |
| *Theme 3: Commendatory Mast* | 24 |
| *Miscellaneous* | 69 |
| 1. Unusual stories, not elsewhere classified | |
| 2. Stories containing strong indication of pathology | |
| 3. Omissions | |

neither of the men get the girl. In the remaining stories there is some attempt at solution whereby *both* of the men win the girl.

A second theme, which occurs in approximately one-third of the stories, revolves around the girl being either the wife or the sweetheart of one of the men and there is no overtly expressed competition over the girl. In these themes more than four-fifths of the stories depict the girl as belonging to the enlisted man and in many of the stories some element of identification with the officer is present.

Theme 3, which occurs in about 20 per cent of the stories, is important clinically, for here we believe we see some of the beginnings of overt anxiety about the subject's rôle in the military and possible disturbances with authority relations. In this theme the officer is severely reprimanding the sailor for some offense or shows some form of strong disapproval.

TABLE 2
PLATE 8: SAILOR AND OFFICER WITH GIRL ($N = 914$)*

| | Frequency | |
|---|---|---|
| *Theme 1: Competition between Officer and Enlisted Man Over Girl* | | 347 |
| 1. Both get the girl | 19 | |
| 2. Girl chooses the officer | 101 | |
| 3. Enlisted man gets the girl | 75 | |
| 4. Enlisted man gets girl—no struggle apparent | 12 | |
| 5. Girl rejects both | 7 | |
| 6. Either or both men reject the girl | 31 | |
| 7. Neither get girl; no reason given | 2 | |
| 8. Don't know the outcome | 100 | |
| *Theme 2: Wife or Sweetheart of Either Man—Competition not so apparent or non-existent* | | 289 |
| 1. Girl belongs to officer | 10 | |
| 2. Girl belongs to sailor | 260 | |
| 3. Girl belongs to officer—is caught out with sailor | 19 | |
| *Theme 3: Officer and Enlisted Man's Relationship the Closest; Woman Secondary or not Mentioned* | | 19 |
| *Theme 4: Officer is Meeting the Couple; in most instances he is showing disapproval of their behavior* | | 188 |
| 1. Generally disapproves | 52 | |
| 2. V.D. warning by officer | 19 | |
| 3. Sexuality main theme | 13 | |
| 4. Guilty conscience | 4 | |
| 5. Officer warns sailor | 41 | |
| 6. Shore patrol officer | 11 | |
| 7. Friendliness and advice by officer | 5 | |
| 8. Military violation—didn't salute, etc. | 22 | |
| 9. Officer doesn't approve of girl—sailor leaves her | 3 | |
| 10. Reprimands girl | 8 | |
| 11. Officer merely greets couple, passes on | 10 | |
| *Theme 5: Girl is Related to One of the Men* | | 37 |
| 1. Sister of officer; officer approves of enlisted man | 7 | |
| 2. Sister of officer; officer doesn't approve of enlisted man | 10 | |
| 3. Commanding Officer's daughter—approved by officer | 4 | |
| 4. Commanding Officer's daughter—not approved by officer | 3 | |
| 5. Commanding Officer's daughter—don't know outcome | 3 | |
| 6. Misc. relative of girl | 3 | |
| 7. Enlisted man and girl are brother and sister | 7 | |
| *Theme 6: Miscellaneous* | | 66 |
| 1. Blanks | 16 | |
| 2. No relation between the three people | 4 | |
| 3. Theme not coherent | 8 | |
| 4. Incomplete | 4 | |
| 5. Clinical material | 12 | |
| 6. Stories not able to classify | 19 | |
| 7. Several themes | 3 | |

*Conflicting totals due to some stories containing two complete themes.

Themes 4 and 5 occur with less frequency (5 per cent or less) but are believed to be even more significant clinically. Here we find stories in which the girl is directly related to the officer or to the enlisted man and thus is not involved sexually in direct competition. Some of these stories possibly reflect strong unresolved disturbances in the oedipal area and sibling rivalry. In another theme, the relationship between the officer and the enlisted man becomes the strongest, that of the girl with either of the men is secondary, or is insignificant, or is completely ignored.

Table 2 presents a categorization of stories given by 914 submarine candidates selected at random from those being processed during 1951-52.

### 3. *Plate 10: The Happy Sailor*

The submarine candidates' stories on this picture can roughly be broken into two groups. The first portrays varying amounts of enthusiasm for a new assignment. Here the predominant theme is that of a recruit either going on "boot" leave or to his first assignment. Another sizeable portion of the candidates give the story of a sailor who is getting out of the Navy. There are various attitudes expressed, from regret to extreme dissatisfaction.

## E. SUMMARY

In searching for suitable techniques for personality assessment in the Naval service, Murray's Thematic Apperception Test (*TAT*) was selected for use in ferreting out factors at a deeper level than could be obtained in a brief interview. Several modifications in the method of administration and in the test stimulus were experimented with. The form now being used consists of 10 pictures, especially drawn for the Navy series. Seven of the pictures used were adapted to the Navy from original Murray pictures, and three new ones were drawn, depicting situations peculiar to the Navy. The technique selected consists of reproducing the pictures approximately three by four inches at the top of sheets of paper which may then be selectively arranged in booklets. The candidates are then asked to write stories about each picture and the stories may then be interpreted according to any number of standard *TAT* systems.

This paper presents a preliminary report on the technique with illustrative themes obtained on the three new pictures by groups of enlisted submarine candidates. The test now is being used experimentally as a part of the basic psychological battery given to all enlisted candidates for the U. S. Naval Submarine School at New London, Connecticut.

## REFERENCES

1. LYON, W. B., MOLISH, H. B., & BRIGGS, D. L. The Cornell index: a comparison of psychiatric "suspects" and "non-suspects." *Armed Forces Med. J.*, 1953, **4**, 977-985.
2. MOLISH, H. B., LYON, W. B., & BRIGGS, D. L. Character structure of adjusted and maladjusted naval recruits as measured by the blacky pictures. *Amer. J. Orthopsychiat.*, (in press).
3. MURRAY, H. A. Explorations in Personality. New York: Oxford, 1938.
4. ———. Thematic Apperception Test Manual. Cambridge: Harvard, 1943.
5. U. S. NAVAL MEDICAL RESEARCH LABORATORY ASSESSMENT STAFF. The assessment program for the submarine forces: studies in progress. Naval Medical Research Laboratory Memorandum Report 52-11, 1952.

*Neuropsychiatric Service*
*U. S. Naval Hospital*
*Oakland 14, California*

Published as a separate and in *The Journal of Psychology*, 1954, **37**, 243-249.

# AN EXPERIMENTAL EVALUATION OF THE RELATIVE EFFECTIVENESS OF INSULIN COMA AND ELECTRONARCOSIS THERAPIES WITH SCHIZOPHRENIC PATIENTS: I. TWO WEEKS AFTER TREATMENT*

*Department of Psychology, University of California*

JAMES C. COLEMAN AND LEON I. HELLMAN[1, 2]

## A. INTRODUCTION

Although a number of studies have been published on the effectiveness of electronarcosis and insulin coma therapies in schizophrenia, the detailed comparison of these two therapies on equivalent patient populations is lacking. The present investigation is an attempt to help fill this gap in psychiatric knowledge. Essentially it explores the comparative effectiveness of *ICT* and *EN* treatments of schizophrenia by means of: (*a*) psychological test changes, (*b*) psychiatric evaluations, and (*c*) the relationship between these two independent evaluations.

## B. PROCEDURE

Sixty male psychotic veterans of World War II were selected for the present study on the basis of the following criteria: (*a*) diagnosis of schizophrenia, paranoid type, (*b*) under 40 years of age, (*c*) psychosis of less than two years duration, (*d*) no prior shock treatment, and (*e*) psychiatric recommendation of shock treatment.

These patients were placed randomly into two treatment groups—*ICT* and *EN*. The equivalence of the two groups was further established by a comparison of the pre-treatment psychological test scores for each group which revealed no significant difference. Two fatalities occurred in the *ECT* treatment group and these cases were excluded from this study, leaving a total of 58 patients who completed the psychological and psychiatric evaluations.

---

*Received in the Editorial Office on November 26, 1953, and published immediately at Provincetown, Massachusetts. Copyright by the Journal Press.

[1] From the Psychological Service, Brentwood Hospital, Veterans Administration Center, and the Department of Psychology, University of California at Los Angeles. Reviewed by the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions of the authors are the result of their own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

[2] The authors wish to acknowledge the coöperation and assistance of Samuel A. Reznick, M.D., in connection with the psychiatric appraisal.

The technique used for the administration of *ICT* was essentially that described by Bond and Shurley (1), and the technique used for the administration of *EN* was essentially that described by Tietz *et al.* (2). In order to exercise better controls, the patients in both experimental groups were placed on the same ward and given similar ward treatment. The diet was the same for both groups, all patients received occupational therapy at the same time with the same therapists, and other ward conditions were equated as far as possible.

All patients were subjected to psychological tests and to psychiatric appraisal prior to treatment and two weeks after the cessation of treatment. Psychological tests provided 36 measures of mental functioning in the following seven areas: new learning, old learning, psychomotor speed, perception, verbal intelligence, performance intelligence, and general intelligence. Psychiatric changes were assessed on the basis of the following rating categories: remission, social improvement, administrative improvement, and no improvement. In no instance did the person doing the psychological testing or making the psychiatric appraisal know the type of treatment.

## C. FINDINGS

### 1. *Psychological Appraisal*

A summary of the psychological test findings is shown in Table 1. The *ICT* group showed significant post treatment gains on 21 of the 36 test variables. Thirteen of the remaining 15 test variables showed gains which were not statistically significant. On no test variable did the *ICT* group show a significant loss on post treatment test scores. In contrast the *EN* group showed no significant post treatment gains in test scores but revealed significant losses for 10 of the test variables. In addition, 16 of the remaining 26 test variables showed losses which were not significant. Generally speaking, the *ICT* group showed significant test changes in a positive direction whereas the *EN* group showed significant test changes in a negative direction.

Although post test changes consistently favored the *ICT* group, these changes were not distributed over all test variables. The question arose, therefore, as to which areas of mental functioning were most subject to change as a result of each treatment. Turning to Table 1, it may be seen that: (*a*) new learning—was favored by *ICT* with six out of eight variables showing significant gains, while *EN* had a negative effect with two of the eight variables showing significant losses; (*b*) old learning—was favored by *ICT* with one out of three test variables showing a significant gain, while *EN* had an adverse effect with one of the three variables showing a significant

TABLE 1

SUMMARY OF PSYCHOLOGICAL TEST CHANGES FOLLOWING *ICT* AND *EN* TREATMENTS

| Area | Mean change *ICT* | significance | Mean change *EN* | significance |
|---|---|---|---|---|
| New Learning | | | | |
| Digit Span Forward | .200 | NS | — .571 | NS |
| Digit Span Backward (Wechsler-Bellevue) | .867 | VS | — .143 | NS |
| Digit Span (Wechsler-Bellevue) | 1.367 | VS | —1.095 | NS |
| Object Memory | | | | |
| Trial 1 | 2.200 | VS | .158 | NS |
| Trial 2 | 1.440 | S | — .632 | S |
| Paragraph Memory (Wechsler-Memory Scale) | | | | |
| Passage 1 | 1.571 | S | .333 | NS |
| Passage 2 | 1.429 | S | .833 | NS |
| Associated Learning (Wechsler-Memory Scale) | .464 | NS | —3.184 | VS |
| Old Learning | | | | |
| Information (Wechsler-Bellevue) | .633 | NS | —2.333 | VS |
| Vocabulary (Wechsler-Bellevue) | .800 | S | — .619 | NS |
| Vocabulary (Shipley-Hartford) | .719 | NS | — .650 | NS |
| Psychomotor Speed | | | | |
| Digit Symbol (Wechsler-Bellevue) | .267 | NS | —1.524 | S |
| Figure Eights | 1.433 | NS | — .368 | NS |
| Color Naming (Wells & Ruesch) | 4.786 | NS | — .650 | NS |
| Fluency | 1.900 | NS | —5.200 | S |
| Perception | | | | |
| "Beta" (Same and opposites from Revised Beta) | | | | |
| 3 minutes | 8.050 | VS | — .278 | NS |
| 3 minutes—errors | .050 | NS | .621 | NS |
| Total time | 39.129 | S | 8.389 | NS |
| Total Errors | 1.226 | S | — .222 | NS |
| Grayson Perceptualization | | | | |
| Passage 1 | 6.414 | VS | —1.650 | NS |
| Passage 2 | 7.621 | VS | 4.050 | NS |
| Intelligence, Verbal | | | | |
| Verbal *IQ* (Wechsler-Bellevue) | 6.700 | VS | —6.667 | S |
| Verbal Age (Shipley-Hartford) | .353 | NS | — .240 | NS |
| Abstraction Age (Shipley-Hartford) | 1.509 | VS | .385 | NS |
| Conceptual Quotient (Shipley-Hartford) | 8.875 | VS | 2.053 | NS |
| Comprehension (Wechsler-Bellevue) | 1.467 | S | — .762 | NS |
| Arithmetic (Wechsler-Bellevue) | .533 | NS | — .810 | NS |
| Similarities (Wechsler-Bellevue) | 1.600 | VS | —1.333 | S |
| Abstraction (Shipley-Hartford) | 5.000 | VS | .385 | NS |
| Intelligence, Performance | | | | |
| Picture Arrangement (Wechsler-Bellevue) | .433 | NS | — .619 | NS |
| Picture Completion (Wechsler-Bellevue) | .800 | NS | —1.095 | NS |
| Object Assembly (Wechsler-Bellevue) | .200 | NS | —1.619 | S |
| Block Design (Wechsler-Bellevue) | 1.033 | VS | — .667 | NS |
| Performance *IQ* (Wechsler-Bellevue) | 2.500 | NS | —7.285 | S |
| Intelligence, General | | | | |
| *IQ* (Wechsler-Bellevue) | 5.067 | VS | —7.333 | S |
| *MA* (Shipley-Hartford) | 1.216 | VS | .110 | NS |

VS = Significance at 1 per cent.
S = Significance at 5 per cent.
NS = Not significant.
+ = Gain.
— = Loss.

loss; (*c*) psychomotor speed—no significant changes were shown in the *ICT* group while two of the four test variables showed significant losses in the *EN* group; (*d*) perception—was favored by *ICT* with five out of six test variables showing significant gains while the *EN* group showed no significant changes; (*e*) verbal intelligence—was favored by *ICT* with six out of the eight test variables showing significant gains, while two of the test variables showed significant losses in the *EN* group; (*f*) performance intelligence—evidenced equivocal results for the *ICT* group with only one of the five test variables showing a significant gain, while two of the variables showed significant loss for the *EN* group; (*g*) general intelligence—was favored by *ICT* with a significant gain for this group while the *EN* group evidenced a significant loss.

In general *ICT* proved superior to *EN* in all areas of mental functioning covered in the present study. With respect to specific areas *ICT* was most effective in producing gains in new learning, perception, verbal intelligence, and general intelligence. *ICT* produced equivocal results in old learning and performance type intelligence tasks. On the other hand *EN* produced significant losses in all test areas except perception.

### 2. *Psychiatric Appraisal*

The psychiatric ratings were made on the basis of the following four categories: (*a*) Remission—return to pre-morbid status, good adjustment, and good insight. (*b*) Social Improvement—return to good social adjustment, but with residuals of illness as manifested in such symptoms as emotional dullness and anxiety. (*c*) Administrative Improvement—lessening of intensity of symptoms, but with fundamental process unchanged. (*d*) Unimproved—no improvement noted. Patients classified in the first two categories were considered by the psychiatrists as sufficiently improved to function adequately outside the hospital. This, of course, was not true for the patients in the other two categories.

Preliminary inspection of the classification of the patients into the four categories noted above failed to reveal any differences in the effects of the two forms of treatment, possibly because of the small number of cases in each category. Therefore the patients were classified into two major groupings of improved and unimproved, with results as indicated in Table 2. The psychiatric appraisal of these two types of treatment revealed no significant difference between them.

TABLE 2
PSYCHIATRIC EVALUATION OF IMPROVEMENT WITH *ICT* AND *EN*

| | Number of patients | | Per cent of |
|---|---|---|---|
| | Improved | Unimproved | patients improved |
| *ICT* | 17 | 11 | 60 |
| *EN* | 16 | 14 | 53 |

### 3. *Relationship between Psychological and Psychiatric Appraisals*

Inspection of correlation coefficients between psychological test changes and psychiatric behavior ratings indicated that there was no consistent relationship between these two methods of appraisal of patients, at least within the time limitations of the present experiment. It is considered possible that followup studies may yield a relationship which is not present in this study.

## D. DISCUSSION

The discussion and evaluation of the present findings can perhaps be most meaningfully made within the framework of what we may term "ideal" criteria for evaluating the relative effectiveness of *ICT* and *EN*. These criteria would essentially include: (*a*) relative effectiveness of each treatment on the total personality of the patients as ascertained by psychological and psychiatric appraisals, (*b*) the permanence or transitoriness of any differential findings as ascertained by a long range study, (*c*) administrative considerations relating to the comparative hazards of each treatment, ease of administration in terms of personnel and time required, expensiveness, and similar factors. Although such an overall evaluation was outside the scope of this investigation, it may serve as a reference point for evaluating our present findings.

Criterion (*a*) was not completely met in the present experiment since the psychological tests reported measured only the cognitive aspects of personality functioning rather than "total personality." Although distorted cognition is typical of schizophrenic patients and treatment changes here are of great significance, the affective and conative aspects of personality are of equal importance and were not covered in the present psychological appraisal. The psychological findings in cognitive areas did, however, unequivocably favor *ICT* over *EN* with statistically significant improvement following *ICT* treatment for the majority of tests and areas of mental functioning covered. In no case did *ICT* result in a significant loss of function. Conversely *EN* treatment resulted in significant losses in the majority of tests and areas covered, and in no case did test performance improve significantly following

*EN* treatment. In this context it is interesting to note that psychiatric appraisals, which were presumably judgments of the "total" personality, failed to reveal significant differences between the two methods of treatment and showed no consistent relation to psychological test findings, thus suggesting the hypothesis that different variables were being appraised by the psychological and psychiatric procedures.

The second criterion, stability of findings, was not applicable in the present study which dealt only with a comparison of the effects noted within two weeks of the cessation of treatments. Since the impairment of intellective processes resulting from electric shock treatment is of temporary duration only, it might be inferred that the same result would hold true for *EN*. If this inference is accurate, the significance of the obtained differences in this experiment might be markedly reduced or even obliterated. A followup study which is under way should be helpful in clarifying this particular problem.

The third criterion (administrative considerations) was found in the present investigation to definitely favor *EN* over *ICT* treatment. There were two fatalities in the *ICT* group as compared to none in the *EN* group and the number of personnel involved and expense made the *ICT* treatment generally less desirable from an administrative viewpoint.

## E. SUMMARY

The present study investigated the comparative effectiveness of *ICT* and *EN* treatments of schizophrenia by means of: (*a*) psychological test changes, (*b*) psychiatric appraisal of improvement, and (*c*) the relationship between these two independent evaluations. The major findings were as follows:

1. Significant improvement in psychological test performance followed *ICT* treatment in most areas of mental functioning measured.
2. Significant losses in psychological test performance followed *EN* treatment in most areas of mental functioning measured.
3. In no case was psychological test performance impaired by *ICT* treatment or improved by *EN* treatment.
4. Improved psychological test performance following *ICT* treatment was most marked in the areas of new learning, old learning, perception, verbal intelligence, and general intelligence.
5. No statistically significant difference was found between the two treatments on the basis of psychiatric evaluations.
6. No significant relationship was found between the psychological and psychiatric indices of changes resulting from either *ICT* or *EN* treatments.

7. The major limitations of the present findings were considered to be (*a*) the restriction of psychological tests to cognitive aspects of personality functioning, and (*b*) the time limitation of the psychological and psychiatric evaluations to a period of two weeks following the cessation of treatment. It is hoped that the followup study currently under way, as well as a future investigation directed toward a broader coverage of personality variables, will help clear up questions relating to these limitations.

## REFERENCES

1. BOND, E. D., & SHURLEY, J. T. Insulin therapy and its future. *Amer. J. Psychiat.*, 1946, **103**, 338-341.
2. TIETZ, E. B. Further experience in electronarcosis. *J. Nerv. & Ment. Dis.*, 1947, **106**, 150-158.

*Department of Psychology*
*University of California*
*Los Angeles 24, California*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 251-256a.

# A MATHEMATICAL DESCRIPTION OF THE RELATIONSHIP BETWEEN FREQUENCY OF RESPONDING (*F*) AND INTERTRIAL INTERVAL IN A MOTIVATIONAL SITUATION*

*Institute of Educational Research, Indiana University*

E. MECH, E. KAPOS, F. HURST, AND D. AUBLE[1]

## A. THE PROBLEM

The current lack of any vigorous attempt toward a systematic description of the variables important to schoolroom motivation presents an arresting situation for the educational psychologist.

Since verbal interaction with pupils appears to be one aspect of the classroom environment which teachers can to an extent manipulate, it seems plausible to inquire into the motivational effectiveness of certain of these verbal behaviors. More precisely, the present paper is concerned with a mathematical description of the relative effectiveness of verbal reinforcement or praise when the reinforcements are administered in both a massed and spaced fashion.

What we are asking is whether organisms perform differentially when motivation (verbal reinforcements) is massed or when it is spaced. Viewed in terms of a classroom situation—if a teacher is going to reinforce a pupil by verbal techniques—does it matter whether the reinforcements are massed or whether they are spaced?

## B. METHOD

In order to obtain data relevant to the foregoing problem the following operations were carried out.

Fourth-grade pupils, 240 in number, were divided into two large groups to receive massed and spaced training. These two large groups were randomly divided into three more subgroups to receive respectively 0, 50, and 100 per cent verbal reinforcement.

The task consisted of routine arithmetical computation for 14 trials. A trial consisted of a four-minute period during which pupils worked routine

---

*Received in the Editorial Office on December 3, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors are indebted to Dr. Nicholas A. Fattu, Director, Institute of Educational Research, Indiana University, for his support.

arithmetical problems. Sufficient problems were provided for each trial so that the most rapid subject was unable to complete the list within the time limit. Ten minutes between each of the 14 trials defined the spacing of trials, while the massing of trials was defined as administering each of 14 trials in immediate succession. The verbal reinforcement was similar to that used by Auble and Mech (1, 2, 3, 4) in previous classroom investigations.

## C. RESULTS

In order to test for possible differences between the massed and spaced training groups during the 14 trials, an analysis of covariance was computed between the pre-test Trial 1 scores and the cumulated scores of Trial 14.

The analysis summarized in Table 1 indicates, at the .01 level of confidence, that the group receiving massed training performed differently from the group receiving spaced training.

TABLE 1
SUMMARY OF THE COVARIANCE ANALYSIS FOR DIFFERENCES IN RESPONSE FREQUENCY DURING 14 TRIALS BETWEEN THE MASSED AND SPACED GROUPS

| Source of variation | Degrees of freedom | Mean square | F |
|---|---|---|---|
| Inter-trial interval | 1 | 104.72 | 11.73** |
| Residual | 233 | 8.93 | |
| Sum | 234 | | |

**Significant at the .01 level of confidence.

Figures 1, 2, and 3 show response frequencies for *S*s receiving 100, 50, and 0 per cent reinforcement respectively, over 14 massed trials. A smooth curve representing an equation was fitted to the data. In view of the pre-test trial being carried out under uniform conditions, only the last 13 trials were considered in the mathematical analysis. Figures 4, 5, and 6 show similar data for three reinforcement groups that received spaced training.

Lewis (5) points out quite accurately that curve fitting is a sort of trial-and-error process. The foregoing cases proved no exception.

It should be noted that the points in each of the graphs were based upon data derived from taking overlapping sequences of three trials each. These sequences were totaled, but not averaged out. This was done to reduce the trial-by-trial variation. The foregoing operations did not in any way obscure the actual performance trends shown in the graphs.

It is well known that logarithmic and exponential functions are closely related. Figures 4, 5, and 6, showing the curves for spaced training, ap-

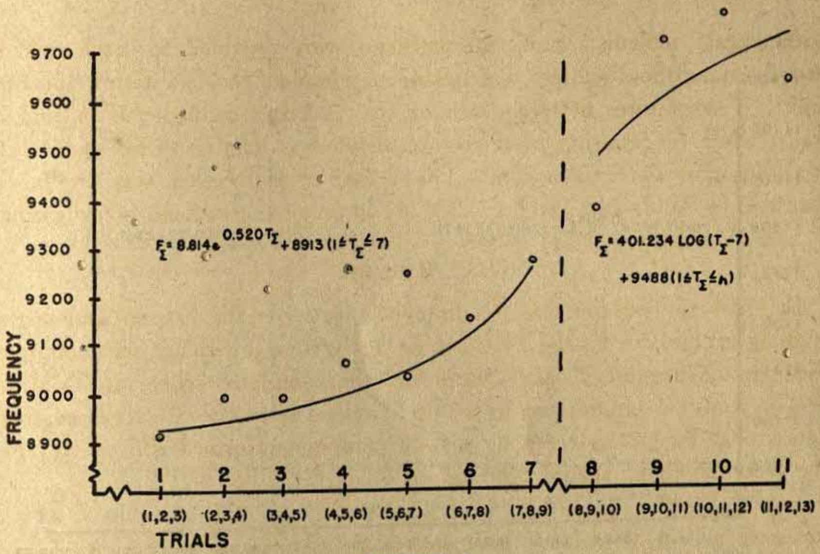


FIGURE 1

Graphic representation of correct responses for fourth-grade *S*s receiving 100 per cent reinforcement over 14 massed trials, together with a smooth curve representing an equation fitted to the data as a function of sequences of overlapping trials (Group $N = 40$).

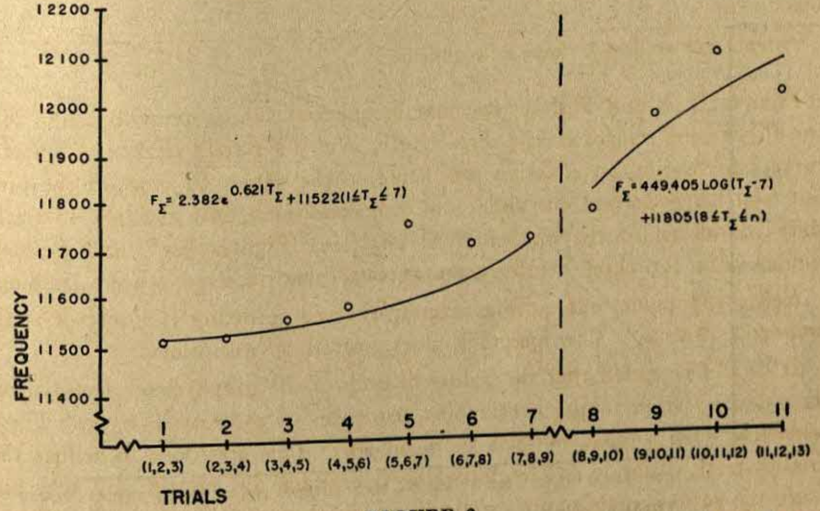


FIGURE 2

Graphic representation of correct responses for fourth-grade *S*s receiving 50 per cent reinforcement over 14 massed trials, together with a smooth curve representing an equation fitted to the data as a function of sequences of overlapping trials (Group $N = 40$).

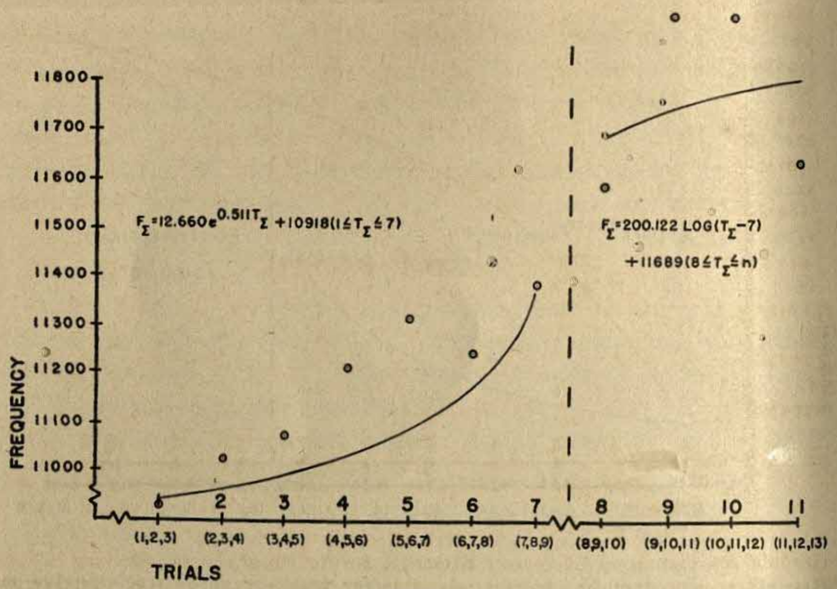


FIGURE 3

Graphic representation of correct responses for fourth-grade *S*s receiving 0 per cent reinforcement over 14 massed trials, together with a smooth curve representing an equation fitted to the data as a function of sequences of overlapping trials (Group $N = 40$).

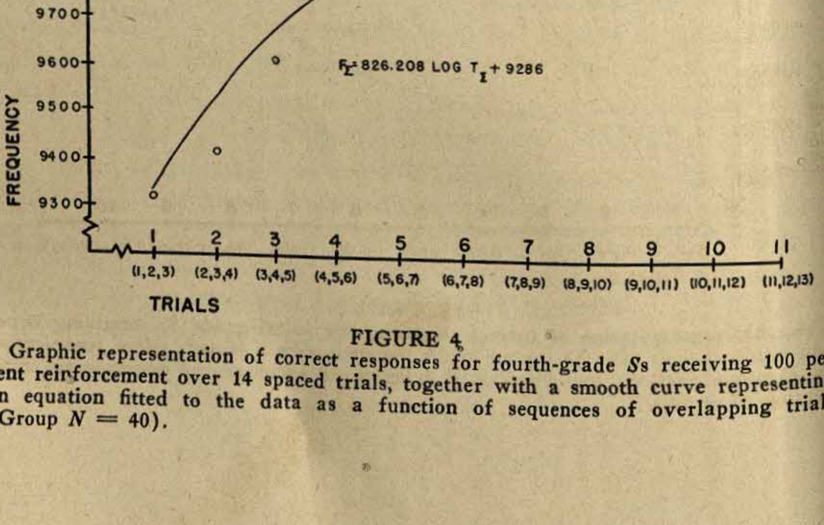


FIGURE 4

Graphic representation of correct responses for fourth-grade *S*s receiving 100 per cent reinforcement over 14 spaced trials, together with a smooth curve representing an equation fitted to the data as a function of sequences of overlapping trials (Group $N = 40$).

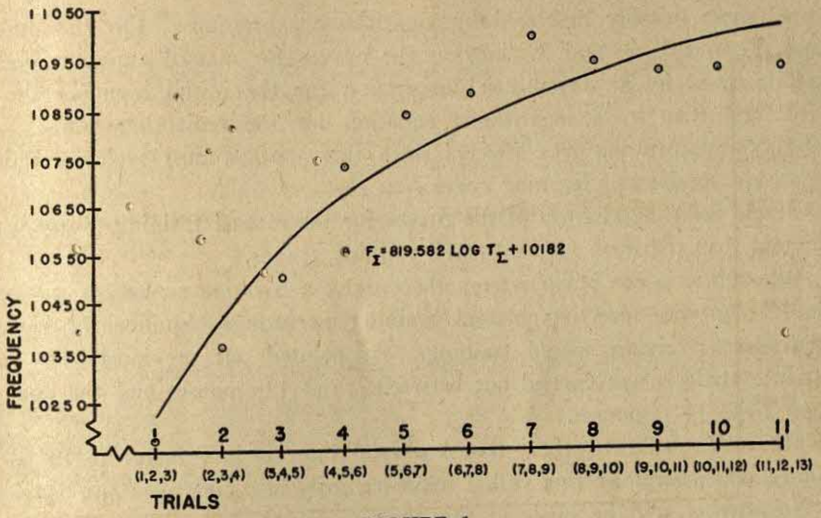


FIGURE 5

Graphic representation of correct responses for fourth-grade *S*s receiving 50 per cent reinforcement over 14 spaced trials, together with a smooth curve representing an equation fitted to the data as a function of sequences of overlapping trials (Group $N = 40$).

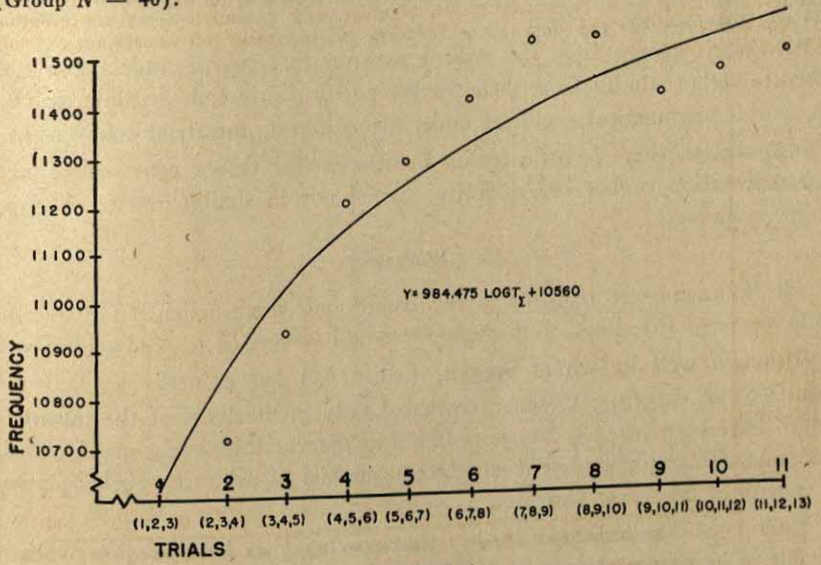


FIGURE 6

Graphic representation of correct responses for fourth-grade *S*s receiving 0 per cent reinforcement over 14 spaced trials, together with a smooth curve representing an equation fitted to the data as a function of sequences of overlapping trials (Group $N = 40$).

peared most suitably described by logarithmic expressions. On the other hand, Figures 1, 2, and 3, showing the curves for massed training, were best described by an exponential expression for the initial sequence of 8 trials, and then by a logarithmic equation for the remaining trials. A logarithmic expression was selected since this appears most typical of the negatively accelerated learning curves.

Clearly the general form of the curves for the massed training group was different than the form for the spaced training groups.

Although it is not obvious from the curves, a covariance analysis showed that the groups receiving massed training performed significantly better than those receiving spaced training. As pointed out previously the covariance analysis was carried out between Trial 1 responses and the cumulated Trial 14 responses.

The curves indicate, then, that intertrial interval is a factor to be considered when administering verbal reinforcements in a classroom situation.

No attempt will be made at this time to construct any "postulates" or "laws" from the scant data at hand. It seems clear, though, that at some future time educational psychologists will be called upon to construct a theory of motivation for schoolroom situations, that can be expressed in quantitative terminology.

All that is known from the present paper is that fourth-grade pupils, exposed to verbal stimuli from their teacher performed in a describable manner on routine arithmetical problems under two differing intertrial conditions.

Nonetheless, there is little reason to believe that theory construction and experimentation cannot be fruitfully carried out in similar restricted classroom situations.

## D. REVISIONS

The following are revisions to the Auble and Mech article (3). Specifically, on page 167, lines 27 through 33 should be revised to read as follows:

Although well known at present, Estes (4a) has pointed out that the results of an experiment can be evaluated only in the light of the measures which have been taken to insure isolation of the variables being manipulated, and that the only completely satisfactory method of achieving such isolation is control of irrelevant variables.

Estes (4a) has indicated further that psychologists have frequently been unable to secure complete control of experimental conditions and have resorted to the use of large numbers of subjects in the hope that these uncontrolled effects would somehow average out.

## REFERENCES

1. AUBLE, D., & MECH, E. V. Quantitative studies of verbal reinforcement in classroom situations: I. Differential reinforcement related to the frequency of error and correct responses. *J. of Psychol.,* 1953, **35**, 307-312.

2. ———. Quantitative studies of verbal reinforcement in classroom situations: II. Occurrence of error and correct responses related to a change in stimulus conditions. *J. Genet. Psychol.,* (in press).

3. ———. Partial verbal reinforcement related to distributed practice in a classroom situation. *J. of Psychol.,* 1953, **36**, 165-186.

4. ———. Response strength in a classroom task related to a "forward" delay in reinforcement. *J. Educ. Psychol.,* (in press).

4a. ESTES, W. K. An experimental study of punishment. *Psychol. Monog.,* 1944, **57**, No. 263.

5. LEWIS, D. Quantitative methods in psychology. Iowa City: Univ. Iowa, 1948.

*Institute of Educational Research*
*Indiana University*
*Bloomington, Indiana*

Published as a separate and in *The Journal of Psychology,* 1954, 37, 257-270.

# PSYCHOLOGICAL PROBLEMS IN MARKSMANSHIP OF INFANTRY TYPE WEAPONS*

*Department of Psychology, Tufts College*

EZRA V. SAUL AND RICHARD S. HIRSCH[1]

## A. THE PROBLEM

The present review is based on an examination of a limited portion of the available literature, and is especially restricted in its coverage of foreign or (military) classified literature. With few exceptions, the marksmanship problems discussed are concerned with direct-fire or line-of-sight weapons such as the pistol, rifle, rocket launcher, etc. However, special reference will be made to those problems arising from the characteristics of indirect-fire weapons, such as are found in the mortar and the grenade launcher.

The following discussion is organized, with appropriate subdivisions, under three headings: Visual Factors in Marksmanship, Motor Factors in Marksmanship, and Marksmanship Training.

## B. VISUAL FACTORS IN MARKSMANSHIP

### 1. *Visual Acuity*

The principal visual function involved in accurate shooting is that of visual acuity. Direct-fire marksmanship assumes the ability of the shooter to perceive the target in a line-of-sight relationship to the sights of the weapon. The minimum acuity level required for proficient shooting depends on the types and distances of the targets and on the variables affecting visual resolution, e.g., contrast, illumination, etc. However, little if any data seem to be available indicating what acuity levels are needed for proficient marksmanship performance at various distances.

In discussing the effects of impaired visual acuity on marksmanship performance, Libasci (14) indicates that the far-sighted individual will have

---

*Received in the Editorial Office on December 4, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This report was prepared in part under Contract Nonr-494(01) between Tufts College and the Special Devices Center, Office of Naval Research. Opinions and conclusions are those of the authors and are not to be construed as necessarily reflecting the view or endorsement of the Navy department.

trouble getting clear definition of either the target or the sights, whereas the near-sighted shooter will have greatest difficulty with target definition at far distances. He also indicates that individuals with presbyopia will have sight pictures which are blurred and indistinct. However, Libasci claims these malfunctions of acuity can, within limits, be modified by corrective lenses.

In the matter of sight and target resolution, Weston (27) and others (18) have pointed out that it is impossible to get clear resolution of both sight and target simultaneously. This is especially true for those situations in which the target is at a considerable distance from the shooter. In this connection Weston indicates that marksmanship performance will be good if the shooter concentrates on maintaining a clearly defined sight alignment and permits target definition to become blurred and indistinct. Army marksmanship manuals (6), however, stress the use of clear target definition (simultaneously with clear sight definition) in shooting. It remains for experimentation to determine whether marksmanship instruction should advocate the maintenance of maximal visual resolution of the sights and/or the target.

### 2. *Ocular Dominance*

Among the other visual functions thought to be of importance in marksmanship is that of ocular dominance. As early as 1915, Doyne (7) concluded that "ocular dominance does not affect marksmanship adversely when rifleman is trained to shoot with both eyes wide open." Experimental verification of the rôle of ocular dominance in marksmanship performance has been attempted by several investigators (2, 13, 22).

Using five different measures of ocular dominance or "eyedness," Banister (2) examined over 1,000 British soldiers and analyzed the relationship between eye-dominance and marksmanship classification (British Army marksmanship categories). The data appear to justify Banister's conclusion ". . . . that the dominant eye is an important factor affecting ability with the rifle, and that the man with the right eye dominant has a considerable advantage, other things being equal, over other men when required to shoot from the right shoulder."

However, several features of the study do not permit unequivocal interpretation of these findings. For one, Banister used groups of British Army personnel from two different commands. The obtained distributions of ocular dominance in these two groups of subjects differed significantly. Secondly, Banister changed, during experimentation, his tests of ocular dominance. The comparability and reliability of his tests are unknown. Thirdly, Banister's subjects had varying amounts and kinds of shooting experience. Finally it was apparent to Banister, as well as others, that the marksmanship categories used as measures of marksmanship ability were of dubious value for experimental research. It is possible that one or several of these factors may have contributed to the obtained results.

Lebensohn (13), in 1942, reported a somewhat similar study on the relationship between eye dominance and marksmanship scores of U. S. Navy recruits. While Banister used subjects with some (or considerable) shooting experience, Lebensohn used men who had had no previous shooting experience. The marksmanship scores were obtained with variations in firing positions and rates of fire. He concluded that the entirely dextral man (right-handed *and* right-eyed) made the best showing when contrasted with all other combinations of handedness and eyedness. Lebensohn indicates that this difference may be accounted for by the mechanical characteristics of the 1903 Springfield rifle used in the tests of marksmanship proficiency. The Springfield rifle is so constructed and operated that entirely dextral men have most efficient access to the weapon's sights and operating bolt. Conversely, individuals who are left-handed and/or left-eyed are at a disadvantage when firing off the right shoulder. This implies an interaction between the physiological and mechanical (weapon) characteristics. Accordingly, the superior performance of dextral shooters noted by these two investigators may be more a function of weapon design than of ocular dominance per se. In short, the weapon design favors the dextral individual.

Simpson and Sommer (22), in 1942, obtained measures of ocular dominance and marksmanship scores for a group of college students. The measures of ocular dominance were determined by using the manoptoscope, the mailing tube, and the peep-hole test. A correlational analysis of the data permitted the authors to conclude that: "It apparently does not make any difference whether the student uses his preferential eye in shooting or the other eye."

Crider (4), in a summary of previous research, indicated that: (a) early investigations were inadequately controlled and ambiguous, (*b*) measures of ocular dominance needed to be perfected, and (*c*) additional research is needed.

The implication for military research is that not only are studies necessary to determine the rôle of ocular dominance in marksmanship, but also investigations may uncover information on shooting the rifle while keeping both eyes open.

### 3. *Other Visual Functions*

The rôle of visual functions in marksmanship performance, other than acuity and ocular dominance, seems to have received slight attention. However, the study by Simpson and Sommer (22) of 180 engineering freshmen obtained measures of lateral imbalance, visual efficiency for right and left eye, eye coördination, distance fusion, ocular dominance, and marksmanship performance. The visual functions were measured with "modified" Betts Telebinocular[2] tests, and the correlation between any of the above functions and marksmanship proficiency, as measured, was not significant. The authors concluded that the ". . . use of Betts visual tests . . . fail to show any important relationship between these visual functions and rifle marksmanship."

However, Drake (8) indicates a correlation of —0.6 between a visual perception test and marksmanship scores in a group of "best," "average," and "poorest" shooters in an R.O.T.C. unit. He also indicates significant correlations between marksmanship performance and acuity of the aiming eye, time scores, and tactual perception scores.

Needless to say, extended research is necessary to clarify the contradictory character of many of these conclusions.

## C. MOTOR FACTORS IN MARKSMANSHIP

### 1. *Steadiness Factors in Marksmanship*

The rôle of muscular coördination and motor abilities in marksmanship performance has occupied the attention of several investigators. In general, there appears to be agreement that motor and postural steadiness correlates positively and to a considerable degree with rifle marksmanship. As early as 1921, Spaeth and Dunham (23) obtained steadiness measures on a modification of the Whipple steadiness testing apparatus and range marksmanship performance for 73 U. S. Army men. A rank order correlation between steadiness and marksmanship scores was of the order 0.61±.106. These authors conclude that ". . . it is clear that steadiness is definitely associated with marksmanship in the case of those men who have had range experience."

In a subsequent attempt to verify the Spaeth and Dunham findings, Seashore and Adams (21) derived intercorrelations between various measures of steadiness and rifle marksmanship scores for 56 members of a basic R.O.T.C. unit as well as for six members of a university rifle team. Use was made of the ataxiometer, the ataxiograph, thrust steadiness, position steadiness, and rifle steadiness in obtaining indices of motor control and steadiness. The authors found correlation coefficients ranging from 0.39 to 0.47 between the measures of steadiness and rifle marksmanship performance. Furthermore, the members of the rifle team had very superior performance on the above battery of steadiness tests.

---

[2] The Betts Telebinocular tests (produced by the Keystone View Company, Meadville, Pa.) are composed of a series of test patterns which are viewed by the subject through a stereoscopic optical system. The modified feature of the Betts tests in the present instance appears to involve the measurement of various visual functions at far distance only. Hence, these investigators did not obtain or utilize other measures of visual performance at close or near distance, which are obtainable with this instrument.

A further study of the motor abilities underlying rifle marksmanship was reported by Humphreys, Buxton, and Taylor (12). This study obtained intercorrelations between postural steadiness, thrust steadiness, rifle steadiness, rifle marksmanship scores, and ratings of shooting experience. The subjects were 47 students in various levels of R.O.T.C. marksmanship training. The authors' conclusion that "marksmanship can be predicted to a considerable extent at least . . . from steadiness . . ." is based on the obtained data. These data indicate a correlation of 0.72 between rifle steadiness and marksmanship. Furthermore, the data demonstrate intercorrelations between the sub-tests of the steadiness battery ranging from 0.39 to 0.89 and are interpreted by the authors as indicating a "steadiness factor" underlying performance on this battery of tests. Finally, these investigators found a correlation of 0.77 between the test battery and marksmanship scores.

These investigations leave little doubt that a real and important relationship exists between steadiness and marksmanship performance. Though the theoretical interpretations of this relationship have varied, it has been suggested that steadiness performance might be used as a predictive measure for selecting men for marksmanship training or as a measure of progress during marksmanship instruction. Here again it appears that extensive systematic research on less restricted populational samples might be especially useful.

Relevant to the steadiness variable in marksmanship training is the finding by Seashore and Koch (20) that postural steadiness performance deteriorates with increasing fatigue, muscular tension, and introspective set. By introspective set is meant the subject's awareness of his own bodily position and movement. It follows from the data of this experiment that shooting sessions should be of relatively brief duration, should not induce excessive muscular tension, and should not induce a set on the part of the shooter to concentrate on his body position or movements. Any of these conditions will tend markedly to decrease the subject's postural steadiness.

### 2. *Tremor and Muscular Adaptation to Gunfire*

Edwards (9) has demonstrated that the amount of finger tremor in response to recordings of battle sounds is significantly different between groups of veterans and non-veterans, combat veterans and non-combat veterans, and hospitalized (neurotic) veterans and non-hospitalized veterans. It appears that extended exposure to sounds of gunfire leads to accentuation of finger tremor. This appears contrary to the expectation that a reduction of finger tremor will occur as a result of repeated and extended exposure to sounds of gunfire.

A more direct attack on the problem of the adaptation of muscular tension responses to gunfire was done by Davis and Van Liere (5). These investigators studied muscular action potentials of subjects in response to gun shots in a highly controlled laboratory situation. They found that the large initial muscular responses to gun shots do not diminish in amplitude but do diminish somewhat in duration. However, the secondary muscle reactions do undergo diminution of both amplitude and duration as a result of continued exposure to gunfire. The investigators conclude that the findings support the hypothesis that large, initial muscle responses resist adaptation.

Though the findings of Edwards (9) and Davis and Van Liere (5) do not have direct relevance to the rôle of finger tremor and muscular adaptation in marksmanship performance, they do suggest the existence of significant relationships between such variables.

## D. MARKSMANSHIP TRAINING

The bulk of the literature on marksmanship training is based on authoritative and expert opinion. Little, if any experimental evidence is available demonstrating the validity of most of the training principles and procedures currently in use. This is not to say that such principles and procedures do not possess some face validity. Rather, little experimental verification has been achieved. Moreover, it is possible that rigorous evaluation might show some practices to be of questionable value.

One of the more systematic approaches to marksmanship training is found in military training manuals. Illustrative of such manuals is the U. S. Army's Field Manual on the U. S. Rifle, Caliber .30, M1 (6). It is indicated in this text that the fundamentals of rifle marksmanship are: (*a*) correct sighting and aiming, (*b*) correct body positions, (c) correct trigger squeeze and breath control, (*d*) correct application of sustained and rapid fire principles, and (*e*) knowledge of sight adjustments.

No more adequate and inclusive division of topics has been discovered in the limited literature surveyed. Moreover, these categories appear to have applicability to a wide variety of weapons. Accordingly, these categories will be utilized in the following discussion.

### 1. *Training in Sighting and Aiming*

Training in correct sighting and aiming in marksmanship is based on the fundamental principle that accurate shooting can occur only when critical relationships are met involving the target, the weapon's sights, and the shooter's sighting eye. Therefore, training in sighting and aiming constitutes a considerable portion of the training time in the preliminary rifle instruction of the U. S. Army recruit. It is in this phase of instruction that the student riflleman is taught the correct "sight picture" and "sight alignment." Instruction involves the use of two-dimensional pictorial material as well as training devices known as "sighting bars." The "sighting bar" permits the student to view representations of the sights of the weapon and the target bull's-eye in the same relationship as is necessary for the accurate use of the weapon. Several types of sighting bars have been described by the U. S. Army and the Canadian Army (6, 26, 29).

Incorporated within the training on sighting and aiming are exercises, using sighting bars and rifles. These exercises permit the student to make repeated judgments as to correct alignment between a fixed rifle or sighting bar and a movable target bull some distance away. The student's judgments are plotted on a white plot-board immediately behind the target bull, and close groupings of judgments are interpreted as indicating a consistent sighting pattern on the part of the student. The objective of this training is not only to produce "shot groupings," which are as small as possible, but also to teach students to obtain sight pictures which leave as much of the target visible as possible.

The training on sighting and aiming so far described presupposes that the student aims at a point target such as the bull in a conventional bull's-eye target. However, several marksmanship experts such as Weston (27) and Reeves (18) have pointed out that it is impossible to maintain maximal resolution of both targets and sights simultaneously. Furthermore, it is Weston's contention that the more basic skill of correctly aligning the front and rear sights of a weapon should be mastered prior to practice on correctly aligning the front and rear sights with that of a point target some distance away. To this end, Weston recommends that training in sighting and aiming include a preliminary period of practice in sighting the weapon at the center of a blank wall or target area, concentrating on correctly aligning the front and rear sights only. Then when such correct sight alignment is mastered, the student proceeds to sighting and aiming at a bull's-eye, permitting the bull to be blurred and indistinct while the sights are clearly resolved and aligned on the indistinct bull's-eye area. This technique, it is claimed, minimizes distraction to correct sight alignment, and tends to reduce the student's tendency to jerk the trigger when he seeks and gets clear target resolution in front of his sights. In contrast to the approach of Weston, Army marksmanship training teaches the student marksman to keep his target clearly defined in his "sight picture."

It is conceivable that the merits of such divergent approaches to sighting and aiming training as well as the determination of the most efficient types of training targets could be evaluated by carefully designed experimentations.

In a recent article, Reeves (19) suggests that the use of colored sights in hand guns may facilitate correct sight alignment. Should research prove this to be true, the principle of differentially colored sights could apply to a variety of weapons wherein correct sight alignment is critical.

A great variety of aids and devices have been used in sighting and aiming training. The sighting bar, described above, is one device which has been widely used and constantly modified (28). It has also been adapted to a variety of weapons, such as the machine gun (32). Another training device, used by the U. S. and British armies, is the "aim corrector" (26). This "aim corrector" is a mirror-like device attached to the weapon, which permits a coach to see and evaluate the "sight picture" obtained by the person sighting the weapon. Still another device is the sight projector (30), which permits the projection of a sight-target image onto a surface at the time the trigger is pulled. An example of this device is the Devereux-Wappler Sight Projector.

In general, the aids and devices described, as utilized in the training of aiming and sighting, appear to have some measure of face validity, but experimental data justifying the use of such aids seems to be unavailable.

### 2. *Training in Firing Positions*

Following training in sighting and aiming, the student usually receives practice in assuming and maintaining bodily positions used in firing the weapon in question. A variety of positions may be used (6, 26, 29) which give varying amounts of comfort or discomfort to the shooter and varying amounts of steadiness to the weapon. The military literature, however, indicates a preference for certain firing positions, and practice in those positions is stressed during the training program.

Periodically, an expert marksman will propose a novel firing position as being more efficient than those in more common use. For example, Stokes (24) has suggested the use of the "sawbuck" position in firing the U. S. Army carbine. He feels that this position, which permits the shooter to sit cross-legged with his arms well over and resting on his thighs, will steady the weapon to a greater degree than other firing positions. Likewise Applegate (1) advocates the "instinctive pointing" position for firing hand guns in combat situations. The characteristic feature of the "instinctive pointing" position is that the weapon is conceived of as an extension of the hand or finger when that extremity is used in pointing toward an object. Applegate feels that such a pointing position is not only accurate, but is already well established in novice shooters. In light of these considerations, he feels that the use of this position would be more effective than other firing positions for pistols and revolvers.

However, neither Stokes nor Applegate present data demonstrating the advantage of the suggested positions over others. It appears that research is needed to determine which firing positions yield the greatest amount of support and steadiness to the weapon and yet are most comfortable to the shooter.

Training in firing positions is usually accomplished by practice exercises in which students rehearse assuming the various positions. Errors are corrected by coaches, who observe their performance. However, in view of the conclusions of Seashore and Koch (20), such practice should not force the student to be excessively preoccupied with his bodily position. If there is any validity in these findings, such preoccupation could bring about a decrease in the student's steadiness and a concomitant decrease in his marksmanship efficiency.

As an aid to getting into and maintaining firing positions, it is common practice (26, 29) to provide weapon rests in the form of sandbags, wooden boxes, etc., for shooters in the earlier stages of training. Also, the use of a pistol sling has been reported (31) as being effective in training shooters to maintain and to correct firing positions with pistols and revolvers.

### 3. *Training in Trigger Squeeze and Breath Control*

Trigger squeeze refers to the finger response which activates the trigger of the weapon. "Good" trigger squeeze is characterized by the slowly increasing finger pressure exerted upon the trigger so that the exact moment of the trigger's release and the resulting explosion is unknown to the shooter. Under such conditions, the shooter is less likely to flinch in anticipation of the explosion; hence, he is less likely to impart any sudden and undesired movement to the weapon at the moment of firing. It would appear that if the shooter flinches after the explosion, the resultant movement of the weapon will probably not affect the accuracy of the shot since the bullet is usually out of the barrel and on its way to the target within milliseconds after the cartridge has exploded Since poor trigger manipulation can induce undesired weapon movement during firing, trigger squeeze is considered by experts to be a critical phase of instruction.

Practice in trigger squeeze is achieved in a number of ways. In one method (6, 26) the student places his finger over that of an expert who goes through the motions of firing a shot. In another method (6) a coach watches and corrects the student for any overt jerking or flinching in manipulating the trigger. A third method (29) involves the balancing of small coins on parts of the weapon, noting if their position is disturbed by the student during the trigger squeeze. Finally, some military literature (6) mentions the use of intermixed live and dummy cartridges in the magazine of the weapon to demonstrate flinching and trigger squeeze. The use of dummy rounds permits the shooter to observe the adequacy of his trigger squeeze without the occurrence of any distracting explosion and recoil resulting from use of a live round. It might be noted, however, that use of dummy rounds limits the verifying knowledge of results provided by live ammunition.

English (10) describes a device used to translate trigger squeeze into a visual demonstration. A rifle was modified to incorporate a rubber bulb within the weapon's stock. The bulb was then connected to a U-tube filled with a colored liquid. The instructor was then able to describe and demonstrate what correct trigger squeeze would do in changing the height and movement of the colored liquid. Following such a demonstration, students would practice trigger squeeze and try to duplicate the expert's performance on the liquid-filled U-tube. In this connection, Ludlum (15) mentions a pressure dial device used by the Nazis to measure and demonstrate correct trigger squeeze. It should be noted that no reliable and objective data are available indicating the effectiveness of these devices.

"Breath control" is taught as an integral part of training in firing positions and trigger squeeze. Appropriate breath control is considered important in that it prevents certain bodily movements which adversely affect steadiness and marksmanship performance. Students are taught to take a breath, release some of it, and then refrain from further breathing until the shot is fired. According to procedure in military manuals, training in breath control includes the correction of breathing errors by a coach who observes the student during practice. The literature does not indicate how the coach may observe and correct any but the grossest type of respiratory movements.

### 4. *Training in Correct Application of Sustained and Rapid Fire*

Training in the correct application of sustained and rapid fire (6) includes practice of the following sequence: (*a*) rapidly assuming a firing position, (*b*) taking aim, and (*c*) rythmically spacing the firing within a given time period. The literature indicates that emphasis is placed on rapidly becoming set to fire and on recovering quickly between shots so that a given quantity of ammunition may be fired within certain time limits. The literature further indicates that students are urged to strive for accuracy as well as speed. Practice in sustained and rapid fire includes the techniques of counting cadence between shots, loading exercises, and taking firing positions rapidly.

### 5. *Training in the Knowledge of Sight Adjustments*

Knowledge of sight adjustments is necessitated by the need to correct for certain constant errors which occur in the firing of weapons. Such constant errors are attributable to target distance, direction and velocity of wind, characteristics of the specific weapon, and the shooter's unique (constant) sighting and aiming behavior. Many weapons have adjustable rear sights which permit the introduction of compensatory corrections for such constant errors. Such sight adjustments permit the shooter to use the same "sight-picture" or sight-target orientation regardless of the operation of certain sources of error.

When adjustable sights are not available on the weapon, the shooter is taught to use "hold-off" on his weapon. That is to say, the shooter is taught to aim his weapon in such a manner as to compensate for an error factor. For example, the rifleman uses "hold-off" by aiming to the right of the target when a strong wind is blowing from the right of the rifleman-target line. In recent revisions of training manuals (6) "hold-off" is not discussed, since most military weapons now have adjustable sights.

Knowledge of sight adjustments is taught students by training them to "zero" the weapon properly (6). This "zeroing" procedure includes the adjustment of the weapon's sights so that the strike of the bullet coincides with the point at which the shooter aimed. The assumption is that the round has been fired under ideal conditions which include the correct sight picture, trigger squeeze, breath control, and the correct body-weapon position.

Students are taught rules which aid them to estimate the correction necessary to compensate for certain environmental errors such as wind, target distance, etc. The manual (6) then specifies that the student will be instructed in sight characteristics and will be given practice in sight adjustment and "zeroing."

The work of Hendrickson and Schroeder (11) is relevant to the problem of training in sight adjustments. These investigators found that an understanding of the theory of light refraction helped air rifle shooters to hit submerged targets more accurately than shooters who were ignorant of such theory. The implication of this study for general marksmanship training is that, subject to verification, knowledge of a general principle concerning sight adjustments would probably facilitate performance in a variety of situations to which the principle is applicable.

Aids and devices, suggested in the training manuals (6) for teaching sight adjustments include models of the sights of the weapons and charts demonstrating the trajectory caused by certain sources of constant error.

### 6. *Other Factors Relevant to Military Marksmanship Training*

The military literature (6, 26) indicates that certain general procedures are utilized in marksmanship training. These procedures include the "coach and pupil" method, the "talking through" method, and the "cadence counting" method. In the "coach and pupil" method the recruit executing the exercise is observed and corrected by another individual acting as a coach. In the "talking through" method the recruit is directed by the instructor as to when and how to execute various phases of an exercise. The "cadence counting" method, which is highly similar to the "talking through" method, is a technique whereby the instructor gives number signals or commands as cues for the recruit's responding throughout a given exercise.

In reference to another aspect of military marksmanship training, Moreau (17) and Brown (3) have described the use of subcaliber weapons training as preparation for performance on larger caliber weapons. The use of small bore guns and devices permit direct and indirect practice on larger bore weapons without having to expend the more costly ammunition of larger caliber weapons.

In general the literature does not provide data indicating the effectiveness of such training procedures and subcaliber devices.

## E. SUMMARY

The present review of the literature on marksmanship, though limited in its coverage, permits a number of summary generalizations:

1. The literature examined reveals few experimental studies. Most articles, particularly those of a military nature, are based on "expert opinion," anecdotal report, or partial description of existing practices in marksmanship training programs.

2. A number of investigations have attempted to ascertain the rôle of various visual functions in marksmanship performance. Though the data of these experiments are somewhat inconclusive and contradictory, they do indicate the need for a number of experimental studies which will establish the minimum visual skills for marksmanship performance. Furthermore, these studies suggest that inappropriate ocular dominance need not be associated with reduced marksmanship proficiency if the weapon is adequately designed.

3. Other experiments have indicated the existence of useful relationships between measurable motor abilities and marksmanship performance. The promising application of these relationships as selection criteria for marksmanship training and/or measures of learning of marksmanship skills, however, requires some additional research.

4. There is apparent need for an extended evaluation of the effectiveness of a number of training methods, aids, procedures, and devices. Within the limits of the literature examined, there were no data validating the many utilized or advocated marksmanship training methods, procedures, and devices. This was particularly true in literature intended for military use.

## REFERENCES

1. APPLEGATE, R. Training combat pistolmen. *Amer. Rifleman,* 1943, **91**, 14-16.
2. BANISTER, H. A study in eye dominance. *Brit. J. Psychol.,* 1935, **26**, 32-48.
3. BROWN, C. A. Teaching the marines to shoot. *Amer. Rifleman,* 1944, **92**, 10-13.
4. CRIDER, D. The importance of the dominant eye. *J. of Psychol.,* 1943, **16**, 145-151.
5. DAVIS, R. C., & VAN LIERE, D. W. Adaptation of muscular tension response to gunfire. *J. Exp. Psychol.,* 1949, **39**, 114-117.

6. Department of the Army, U. S. Rifle, Caliber .30, M1 Field Manual 23-5/T039A-5AC-11. Washington, D. C.: U. S. Government Printing Office, 1951.
7. Doyne, R. W. Eye in sport shooting. *Ophthalmoscope,* 1915, **13**, 119-126.
8. Drake, C. A. Marksmanship: Aptitude or skill? *Infantry J.,* 1941, **49**, 48-50.
9. Edwards, A. S. Finger tremor and battle sounds. *J. Abn. & Soc. Psychol.,* 1948, **43**, 396-399.
10. English, H. B. How psychology can facilitate military training: A concrete example. *J. Appl. Psychol.,* 1942, **26**, 3-7.
11. Hendrickson, G., & Schroeder, W. H. Transfer of training in learning to hit a submerged target. *J. Educ. Psychol.,* 1941, **32**, 205-213.
12. Humphreys, L. G., Buxton, C. E., & Taylor, H. K. Steadiness and rifle marksmanship. *J. Appl. Psychol.,* 1936, **20**, 680-688.
13. Lebensohn, J. E. Ocular dominance and marksmanship. *Nav. Med. Bull.,* 1942, **40**, 590-594.
14. Libasci, A. M. Eyesight (conclusion). *Amer. Rifleman,* 1950, **98**, 24-26.
15. Ludlum, S. Nazi training. *Amer. Rifleman,* 1942, **90**, 12-14.
16. Moore, C. H. Training device for machine gunners. *Amer. Rifleman,* 1944, **92**, 11.
17. Moreau, J. W. Midshipman marksmanship. *Amer. Rifleman,* 1944, **92**, 10-11.
18. Reeves, H. Here's how. *Amer. Rifleman,* 1944, **92**, 20-24.
19. ———. The international handgun game. *Amer. Rifleman,* 1951, **99**, 44-47.
20. Seashore, H. G., & Koch, C. Postural steadiness under conditions of muscular tension and fatigue. *Psychol. Rec.,* 1938, **2**, 319-332.
21. Seashore, R. H., & Adams, R. D. The measurement of steadiness: A new apparatus and results in marksmanship. *Science,* 1933, **78**, 285-287.
22. Simpson, R. G., & Sommer, R. C. Certain visual functions as related to rifle marksmanship. *Sch. & Soc.,* 1942, **55**, 677-679.
23. Spaeth, R. A., & Dunham, G. C. Correlation between motor control and rifle shooting. *Amer. J. Physiol.,* 1921, **56**, 249-256.
24. Stokes, W. R. The carbine fires for record. *Amer. Rifleman,* 1943, **91**, 18-20.
25. ———. The best heritage. *Amer. Rifleman,* 1944, **92**, 9-10.
26. Thornton, G. E., & Walters, H. de L. Aids to Weapon Training. Great Britain: Aldershot, Gale & Polden, 1942. (Pp. 9-11.)
27. Weston, P. Target Shooting Today. New York: Greenberg, 1950.
28. An improved sighting bar. *Amer. Rifleman,* 1944, **92**, 38.
29. Shoot to Live: Presenting the Johnson method of musketry coaching as adopted by the Canadian Army. Prepared under the direction of the Chief of the General Staff. Ottawa: Cloutien, 1945.
30. Sight projector aids class training. *Amer. Rifleman,* 1944, **92**, 35.

*Institute for Applied Experimental Psychology*
*Tufts College*
*Medford 55, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 271-277.

# ANXIETY AND SOCIAL SENSITIVITY*

*Department of Psychology, VA Hospital, Lebanon, Pennsylvania*

LEO SCHNEIDERMAN

## A. THE SOCIAL AND THE AUTISTIC

Recently, Mowrer (2) has called to our attention the possible rôle of repressed moral obligations in the etiology of neuroses. He has emphasized, as well, the social usefulness of anxiety in the sense that it signalizes that a moral failure has occurred. Implicit in Mowrer's argument is the assumption, we believe, that a moral failure occurs against the background of a uniformly moral *reality.* If we are correct in attributing this supposition to Professor Mowrer, we might proceed to examine its validity, and in addition, endeavor to trace the theoretical consequences of his view of the potentially adaptive function of anxiety. It might be asked, at the very outset, what we mean by psychological "realism" in a broad sense, and then go on to a closer examination of morality as a psychological datum.

Although our approach may seem circuitous, we might first ask if we are correct in taking the view that a social response, for example, is more "realistic" than an autistic response. To be sure, one is no less "real" than the other in the sense that both occur as visible responses to a situation. Yet, when we speak of a social response, we have in mind one that is appropriate to a consensually validated frame of reference, consisting of linguistic and other attributes. In the case of autism, we are confronted with an activity that may be understood only indirectly, and which seems to be the expression of an idiosyncratic attitude. It is a common practice for clinicians, as well as people in general, to view unfavorably the autistic behavior of another person, and to consider a social response as more desirable, and as more nearly "normal." This preference, as we shall try to point out later on, is rooted to some extent in necessity. Though we may regard both the social and the autistic response as different forms of learned behavior, we shall try to make it clear that the former is a general phenomenon, whereas the latter occurs only under special circumstances.

---

*Received in the Editorial Office on December 5, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

Let us suppose, then, that we are dealing with two sets of responses that have been reinforced in the past, with resulting need reduction. The need-reducing function of a social response is observable in those cases where the individual responds to his presumed knowledge of the other's expectations, rather than in strict accord with his own feelings. By being sensitive to the will of another person, and by acting upon that sensitivity, we arrive, or hope to arrive at a relationship in which the other person's potentially hostile actions are averted. It is possible to speak of a reduction of fear as the reinforcing agent that enables one to learn to make social responses.

On the other hand, in those instances where the individual defines a social situation principally in terms of his own feelings, and ignores the individuality of the other person, we may speak of a lack of social sensitivity. Behaviorally, such insensitivity may take the general form of willfulness, expressed either assertively, as in actively assigning to the other person a rôle of one's own making; or passively, as in leaving the rôle of the other completely undefined. Here, too, we are dealing with an instance of need reduction. By ignoring the other person and his needs, and focusing entirely upon our own inclinations, we strive to establish a relationship in which the self-rôle is enabled to retain maximum stability. As long as the variable impact of other personalities is avoided, it is possible to live in a subjective world that is self consistent because it is unchanging. Reduction of ambiguity with a resulting increase in security (it may be a false security) may be regarded, then, as the fixing agent in the acquisition of self-oriented, or autistic responses. The point we wish to make is simply that self-oriented behavior, though it might be quite inappropriate a good deal of the time, nevertheless may be learned because of its need-reducing character. We might go further and assert that autistic behavior even may be *effective* in an objective sense in certain social situations, depending on the underlying power relationship that exists between individuals. For example, a person who is feared, or with whom we strive to identify ourselves because we view him as some sort of authority figure, may successfully carry off his autism, and even make it the cornerstone of his powerful position by learning to manipulate other persons in terms of his own needs. His sheer inaccessibility thus may become the instrumentality of a superior "adjustment" at the expense of others.

The foregoing observations lead to the inference that under certain conditions autism may be more "realistic" than social behavior. From a pragmatic standpoint, one might say that it *works* occasionally. In general, however, it is plain that a person who responds too often to situations without regard to the wishes of others, will be ill adapted to deal flexibly with the demands of social living. We see now that social responses may be assigned a general preeminence, even though it is necessary to recognize more than one kind of psychological "realism." In the section that follows, we shall treat of two kinds of "realism" that we believe to be generically related to the terms we have used above. We shall try to outline the conditions under which the social response—which we may define as being based on one's own wishes, as well as the presumed wishes of other people—may give rise to consequences of a highly adaptive kind, but with an unfortunate outcome for morality, as it is usually conceived.

## B. ANXIETY

From our present standpoint we are led to make two observations about anxiety. First, as we shall try to demonstrate, anxiety is indeed socially useful, as Mowrer has pointed out, but in a new guise, of which the recognition compels us to abandon the concept of an internal and more or less immutable moral standard. Second, the occurrence of a moral failure cannot be viewed as a basis for the emergence of anxiety. Instead, we are prepared to argue that both moral *and* "immoral" actions occur in response to direct social pressures, and are productive of anxiety only when they violate the prescribed definition of the context in which they occur. That is, anxiety may occur as readily in a person who behaves morally in an "immoral" social climate, as in one who acts unjustly in an atmosphere of general good will. With this possibility in mind, we shall propose at a later point, a definition of neurosis in which the super-ego, conceived as the carrier of fixed moral imperatives, has no place.

But what of those countless persons who suffer from neurotic anxiety because of a conflict between opposing wishes? Are we not dealing here with the effects of a strong and inflexible conscience? It may be maintained with as much reason, perhaps, that the sufferer is possessed with a fear of direct retaliation or disapproval by others, rather than an obsessive dread of moral commandments that he has been taught in the past. It is not anxiety in the sense of a vague fear of something unknown, or a reaction that has been detached from its proper object, that is experienced by the sufferer; it is the realization that he is in danger from his fellow men. In a community where it is permissible to commit murder, let us say, in the name of self defense, or "honor," or for some other reason, it may be supposed that anxiety will not be experienced by the murderer.

Freud (1), in his essay, *Thoughts for the Times on War and Death,* has expressed this idea with great clarity, but has left the impression that only in critical times is there likely to occur a suspension of moral imperatives. A less sanguine view might be that in times of peace as well as war society provides a mandate for aggressive actions. It is not enough to say, in this regard, that "dangerous" impulses are so defined precisely because they are anti-social. We must proceed instead to a closer examination of what society encourages and what it enjoins. We may discover that society does not speak out in a clear, commanding voice concerning the lawfulness of expressing impulses that are directly derivative of hostile wishes. We refer here to such unambivalent forms of behavior as verbal aggression, snobbishness, indifference to the feelings of others, pursuit of power, and many similar behavioral forms that not only are widely sanctioned, but actually are considered the *only* appropriate response in many circumstances. The fate of "benign" behavioral forms such as support, permissiveness, empathy, conciliation, etc., that might be thought to stand in an antithetical relation to those above mentioned, is not hard to imagine. The former may take on the character of taboos, and *their* expression may become dangerous. The view that neurotic anxiety develops because we are forced to repress our hostile wishes is incomplete insofar as it does not comprehend the indifferent, and even *facilitative* attitude of society toward various forms of aggression.

We may now ask if neurotic anxiety indeed develops as a consequence of not being permitted to act as one might wish. We should regard it as a reasonable hypothesis that the very opposite is true: that most men are not anxious in a neurotic sense precisely because society permits them to act aggressively in a multitude of directly satisfying ways. It may be opposed to this claim that it makes light of the facts of neurosis. Rather than seeking to abolish the fact of neurotic anxiety, with which we are all familiar, we are concerned with reversing the traditional balance (in the clinician's mind) between neurosis and character disorder. The latter designation no longer need be applied only to psychopaths. It may be expanded to include vast numbers of "respectable" people who are at peace with their own bad intentions, and feel free to act upon them. Similarly, the incidence of true neurotic disturbance needs to be minimized. Possibly, the relation of the neurotic person to the disordered character is that of an exception to the general rule. To be neurotic, one must have an appreciation of unpracticed—including benign—behavioral forms and motivations, and at the same time be in conflict over one's inability to act as one might wish, i.e., contrary to social usage. It might be said that the person who is anxious in a neurotic sense has not repressed his impulses at all, but instead has been impelled to act in keeping with them, at the same time that he is aware of their antitheses. Like the respectable disordered character, the neurotic generally acts the way he is expected to act by his associates, but unlike the former, he is not happy about his actions.

The question of neurotic depression presents a somewhat different problem than that of anxiety. If, as we have suggested, the incorporation of a set of consistent and exclusive moral values is a rare occurrence, why do people become depressed and filled with guilt? Under such circumstances, is it not proper to infer a fear of punishment, since the depressed person himself reminds us of his unworthiness, and begs to be punished.

We often may observe still another component, in addition to the supposedly ambivalent character of neurotic depression, and this is nothing more than self-pity. It is possible to think of the melancholic as a person who feels that despite his essential goodness, something valued has been taken away from him unjustly, or some deserved object has been withheld arbitrarily. It is as if he has carried out a set of previously reinforced actions, only to find at last that the expected reward was to be denied. The source of distress is to be found not in the individual's feeling that he has been unworthy of his proper standards, but in his sensitivity to the implied rejection of him by his fellow men (or by a personified fate), even though he has been deserving of something better.

Neurotic anxiety and depression both may now be seen as the direct result of an interpersonal process in which the sufferer feels that he has lost the good will of his fellows in some way. This formulation implies that neurotic conflicts are not entirely intra-psychic, but involve as well the individual's relation to the presumed reality of his social setting. The person who does not experience neurotic anxiety or depression, and who is also free of other neurotic symptoms, we may suppose to be no less sensitive to other persons and their attitudes toward him. In his case, however, he has no reason to feel estranged from others. The development of such freedom from conflict with the social environment becomes comprehensible if we take a second look at moral "reality" as it exists for the growing child.

As we know, moral percepts are communicated to the child by both word and example. The assumption often has been made by those who think in terms of an internalized conscience that the reality of the word corresponds to the reality of the example. For example, if a child is taught verbally to treat all persons as equals, it is expected that eventually he will learn this lesson, and come to act upon it. The fact that his parents, and the people around him act in quite a different way is ignored as a likely source of confusion for the young learner. If we acknowledge at this point what already is well known—that in most societies there is, and probably always has been a great discrepancy between principle and practice—we cannot help but see that verbal rules frequently are not reinforced by behavioral facts. It is well to ask, then, how learning can be thought to take place in the absence of realistic incentives. Returning to our example above, it is obvious that the child who is so unobserving of his parental behavior as to actually respond to people as if there were no differences between them, soon will be reminded by his parents of numerous qualifications and points of expediency that constitute *their* social "reality," and upon which they base their actions. As a consequence, we may expect that children will be unable to learn to act upon a symbolic "reality" of moral rules that has no counterpart in social "reality."

The moral "reality" of the child, then, is anchored in social expectations. It may be predicted, therefore, that no set of expectations, no matter how consistent or how compelling in an abstract ethical sense, will long be able to guide an individual conscience-wise, without frequent reinforcements in his day to day experience. If we may speak of a conscience at all, it is as a *process* that must be renewed throughout the lifetime of the individual by means of his social sensitivity. The view that anxiety is the consequence of a moral failure must be reinterpreted as referring to a *social* failure. Our proposal does not seek to reduce all morality to a vulgar expediency, but is aimed at calling attention to the "reality" of the immediate social scene, in contrast to that ambiguous and temporally remote "reality" which the super-ego has been thought to represent. If we may be allowed to return to the analogy employed at the beginning of our paper, we might conclude that the person who overcomes his social sensitivity, and succeeds in acting upon moral imperatives that are "real" for him, will be responding autistically, in effect. That such behavior may be "realistic," i.e., receive reinforcement in the form of favorable consequences for the individual in certain situations, is not denied, however.

We are in general agreement with Mowrer's standpoint that emotional disorder needs to be understood in terms of moral issues. Our view differs from his, however, in assigning to anxiety a situational basis, and similarly, in emphasizing the day to day social definition of conscience.

## C. Summary

We have been concerned with showing that anxiety, including neurotic anxiety, is better understood as a reaction to direct social pressure, than as a consequence of internal conflict. By ascribing to social sensitivity a key rôle in the regulation of behavior, it has been our intention to translate the concept of anxiety into concrete terms, and to account for the socialization process without recourse to the operations of an idealized super-ego. Specifically, we have tried to make clear, from a psychological standpoint, the difference between an oral tradition of moral imperatives, and the actuality of social conduct. It has been hypothesized that only those attitudes and behavioral forms will be learned that are rewarded within an immediate social context. The moral imperatives that have currency within a pluralistic value system cannot become part of a permanent conscience as long as they exist mainly as unrewarded ideals.

## References

1. Freud, S. Collected Papers. (Vol. 4.) London: Hogarth Press, 1950.
2. Mowrer, O. H. Learning Theory and Personality Dynamics. New York: Ronald Press, 1950.

*Department of Psychology*
*VA Hospital*
*Lebanon, Pennsylvania*

Published as a separate and in *The Journal of Psychology*, 1954, 37, 279-289.

# PROBABILITY PREDICTIONS BETWEEN AND WITHIN SERIES*[1]

*Department of Psychology, University of Wisconsin*

BRUCE M. ROSS

## A. INTRODUCTION

Many ordinary situations involve not only choices in which there is at best but a probability of success, but the selfsame situation also requires a prior choice of the particular set of probabilities to be considered. For example, one cannot draw a colored ball even from the textbook urn of probability theory until the urn itself has been selected. A simple probability situation would involve only two colors and two urns. Two decks of white cards were made to do in the present experiment for the urns and either the symbol *X* or *O* stamped on the underside of each card was the counterpart of differently colored balls. College student *S*s chose a deck, predicted the symbol on the top card, and then turned the card over to determine if the prediction were correct. The individual *S* was free to alter either his deck choice or symbol prediction on any of the 300 consecutive trials he performed.

The analogy is not complete. The ball selector has a strong notion that his cumulative selections are dependent on the proportionality factor. The symbol predictor may believe he can detect a sequential pattern independent of proportionality. But how will *S*s react if proportionality *is* the sole patterning of the symbol sequence situation? The present investigation attempted to discover the answer by analyzing both deck (between-urn) and symbol (within-urn) preferences in a free choice situation.

In Experiment *A* one of the two 300-card decks in competition for preference was always required to have a symbol proportionality of 50-50 per cent. In other respects also this deck was made as "random" as possible. Competing decks similarly had only stochastic convergence, but they possessed

---

*Received in the Editorial Office on December 11, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is an adaptation of a thesis submitted to the Faculty of the Graduate School of the University of Wisconsin in partial fulfillment of the requirements for the Ph.D. degree.

The complete thesis is on file in the University of Wisconsin library at Madison. The author is indebted to his major professor, Dr. David A. Grant.

asymmetrical symbol proportionalities of 60-40, 70-30, 80-20, 90-10, and 95-05 per cents. For any of these five conditions an optimal solution was the continuous prediction of the most frequently occurring symbol. In addition a 50-50 versus 50-50 per cent group was run as a control condition. Conclusions were reached, in regard to the proportionality dimension, as to how constrained a stochastic series must be before it is empirically detectable either by symbol or series preference. Where significant differences from chance are first observed can be viewed as the threshold of "patterning" or, conversely, as the borderline of "randomness."

In Experiment *B* four decks were used. Three decks had a symbol frequency of 50-50 per cent and the other deck possessed a symbol frequency of either 70-30, 80-20, or 90-10 per cent. Conclusions were drawn concerning the effect of increasing stimulus complexity on deck and symbol preference.

In Experiment *C* the series constraint applied might be termed a "balancing" restriction. One deck (B10) was required to have a 50-50 per cent symbol frequency on exactly every tenth card with no consecutive symbol runs longer than five symbols. The deck for the other condition (B4) was required to have a 50-50 per cent symbol frequency on exactly every fourth card with no consecutive symbol runs greater than two symbols in length. Each of these balanced decks was separately put in competition with the 50-50 per cent decks used in Experiments *A* and *B* which were required to have an exact 50-50 per cent frequency only on exactly every hundredth card. Exact 50-50 per cent "balancing" could, however, occur by chance at other points in these series.

It can be observed that this sort of short term restriction on "randomness" is a reflection of the patterning often used in experimental administration of stimuli or reinforcement schedules. It was the purpose of Experiment *C* to determine whether there was a point at which this sort of "effective" randomization leads to sufficient patterning to be detectable.

## B. METHOD

### 1. *Experimental Design*

The experiment was designed with eight conditions utilizing two-deck comparisons and three conditions utilizing four-deck comparisons. Eighteen *S*s were given each procedure making a total of 198 *S*s. In addition the results from 12 *S*s were discarded without tabulation because they followed directions inaccurately. The experiment as a whole can be viewed as involv-

TABLE 1
THE EXPERIMENTAL CONDITIONS

| A. Two-deck Symbol Frequency Conditions |
|---|
| 50%-50% vs. 50%-50% |
| 60%-40% vs. 50%-50% |
| 70%-30% vs. 50%-50% |
| 80%-20% vs. 50%-50% |
| 90%-10% vs. 50%-50% |
| 95%-05% vs. 50%-50% |
| B. Four-deck Symbol Frequency Conditions |
| 70%-30% vs. 3(50%-50%) |
| 80%-20% vs. 3(50%-50%) |
| 90%-10% vs. 3(50%-50%) |
| C. Balancing Restriction Conditions |
| B10 vs. 50%-50% |
| B4 vs. 50%-50% |

ing three sub-experiments or three different sorts of stimulus complexity comparisons unified by the use of the same stimulus series as a common component or standard stimulus for all conditions. The conditions that were investigated are summarized in Table 1.

### 2. *Stimuli Presentation*

Each two-choice symbol series consisted of a sequence of 300 3x5 cards stamped either with the symbol *X* or *O*. The cards were placed symbol side down on a Shannon arch file. The two or four files used were placed in front of the individual *S* with a card designating the left file as *1* and the right file as *2;* in the case of four files, reading from left to right—*1, 2, 3, 4.* The position of the decks was randomized within all conditions. The series used were not selected randomly, but were modified in every instance after selection from a table of random numbers (1) to conform to the symbol frequency desired and to the mean number of runs for the constraint chosen.[2] This mean number of runs was obtained by using the formula for mean number of runs given by the Wald-Wolfowitz run test (4).

For the most frequently used series, the 50-50 per cent sequence, three separate series were obtained. For each experimental condition a system of staggering the starting points of both series (e.g., the fiftieth card of the original deck was made the top card and the forty-ninth card the bottom card in a new deck order) and alternating the competitive 50-50 per cent series was followed. Staggering deck starting points obviated possible objections that any response series obtained might be representative only of a single, specific, ordered sequence. For all series the run restrictions with which the original series were constructed remained unaltered.

---

[2] A run is defined as any uninterrupted sequence of the same symbol. There may be runs of only one symbol.

### 3. *Subjects and Directions*

The *S*s were from elementary psychology classes at the University of Wisconsin. Approximately one-third of the total *S*s were men and two-thirds women. Each *S* participated in only one condition. Experimental conditions were administered to groups ranging in size from one to four. No *S* in any group received the same stimulus series as other *S*s in the group. All the two-deck conditions were completed before the four-deck conditions were administered. The *S*s filled out IBM True-False answer sheets for the two-deck conditions and IBM Multiple Choice answer sheets for the four-deck conditions. Cardboard masks for covering all but one of the answer sheet options and soft pencils were given the *S*s. The actual symbol *X* or *O* was written on the answer space.

The oral directions given the *S*s follow. The words in parentheses were insertions for the four-deck conditions.

> In front of each of you are two (four) card files. The cards on these files are face down. The front of each card is marked with an *X* or an *O* just as these cards are. (Illustrate). Your task is to try and make as many correct predictions as you can as to which symbol will come next. You do this by first deciding which deck you will guess the top card in. Then in the column corresponding to the deck number mark an *X* or an *O*.
>
> After you have made your prediction, turn the top card over in the deck you have chosen so that you can see whether your prediction has been correct. Then make another prediction in the next space down on the answer sheet and turn up the new top card in the deck you have made your new prediction in. Continue on in this manner at your own speed. You make only one prediction in any one row just as you would in a Multiple-Choice test. Each of you is to cover his answer sheet with a mask so that only one row is exposed at a time. Regardless of what the card you turn up may be and what your previous guess may have been you are free to switch decks or change the symbol guessed at any time. Remember though, your task is to get as many correct predictions as you possibly can. Now I'll illustrate sample possibilities in filling out an answer sheet. (Illustrate.) Are there any questions as to procedure?
>
> Admonitions against cheating were then given the *S*s covering such points as: turning only one card at a time, not looking at, or touching

> the cards before predicting, and keeping the mask over the answer sheet.
>
> Now, as to the purpose of this experiment. To each group of *S*s I give different clues which together with the information gained by observing the deck aids in making correct predictions. In this present group you will all be given the same information with which to try and improve the number of your correction predictions, but each of you has a different pair (group) of decks.
>
> Here are the clues, so listen carefully.
>
> 1. You will be stopped before you can use up all the cards in either (any of the) deck(s).
>
> 2. Improving your predictions may not mean getting every single prediction correct. That is, you may show some improvement even though you continue to make some errors.
>
> 3. One deck is easier than the other(s) for making correct predictions as to which symbol will turn up on the top card.
>
> When you have completed this part of the experiment you will be given a questionnaire concerned with your performance and your general prediction ability.

All *S*s were given the same directions. The instructions given concerning the alleged purpose of the experiment were designed to encourage purposeful behavior over the long series of 300 trials by making the task meaningful to the *S* in spite of the fact no "totally correct" solution was ever possible.

## C. Results

### 1. *Symbol Predictions*

Symbol predictions were analyzed by 60-trial blocks. For the non-50-50 per cent series correlations were obtained for individual *S*s between the number of symbol predictions and the number of symbol occurrences of the preponderant symbol in the 60-trial block. These product-moment correlations for the first and last (fifth) trial blocks give a partial answer to the questions as to how strongly responses are initially guided by proportionality and whether subsequent prediction alters the proportionality influence. Correlation coefficients are tabled in Table 2 together with the number of *S*s on which they were computed.

The correlation for both the first and fifth trial block are uniformly high (.86 to .99). There is no discernible practice trend despite the fact 180 predictions intervened between Blocks 1 and 5. These high correlations are striking because the *S* was free at any time to return to the 50-50 per cent alternative deck. Since some *S*s never predicted in the non-50-50 per cent deck the number of *S*s predicting in each block varies. Correlations for the

TABLE 2
CORRELATION COEFFICIENTS BETWEEN SYMBOL OCCURRENCE AND PREDICTION OF MOST FREQUENTLY OCCURRING SYMBOL

| Two-deck Conditions | Block 1 N | Block 1 r | Block 5 N | Block 5 r |
|---|---|---|---|---|
| 60%-40% | 17 | .89 | 14 | .88 |
| 70%-30% | 17 | .87 | 15 | .94 |
| 80%-20% | 18 | .97 | 16 | .90 |
| 90%-10% | 18 | .96 | 15 | .98 |
| 95%-05% | 18 | .97 | 17 | .86 |
| Four-deck Conditions | | | | |
| 70%-30% | 18 | .96 | 13 | .98 |
| 80%-20% | 16 | .93 | 14 | .96 |
| 90%-10% | 16 | .97 | 16 | .99 |

50-50 per cent deck predictions are not shown, but they were of the same order as those for the non-50-50 per cent decks.

Two previous studies (2, 3), using fewer trials and fewer conditions, have concluded that there is a strong tendency for the selection of alternatives in stochastic two-choice series at the approximate frequency at which they occur. This is true in the present study even though an easily possible optimal solution would involve a continuous choice of the most frequently occurring symbol. And it is true even though the directions seek to encourage "problem solving" responses as opposed to directions specifying that "randomness" is inherent in the task, a common characteristic of the two cited studies (2, 3).

### 2. *Deck Choice*

The deck choice results of Experiments *A, B,* and *C* are presented in Figures 1, 2, and 3 respectively. Per cent deck choice and number of deck choices are graphed against successive 60-trial blocks.

It is evident from Figure 1 that deck choice preferences do not scale according to symbol proportionality. The exception is Block II (Trials 61-120) the only block in this experiment to yield a significant *F*. (A "between-within" analysis gave $F = 8.50$, significant at the .01 level with 4 and 85 *d.f.*) Omitting the unaccountably depressed 60-40 per cent condition a significant *F* was still found. ($F = 14.25$, significant at the .01 level with 3 and 68 *d.f.*) The 90-10 per cent choices progressively decreased after Block II while the 95-05 per cent deck choice selections continued at a high level, suggesting that the infrequent appearance of a minority symbol may enhance preference by focusing interest on this "long shot." However, the changes in deck choice, as previously pointed out by the symbol prediction

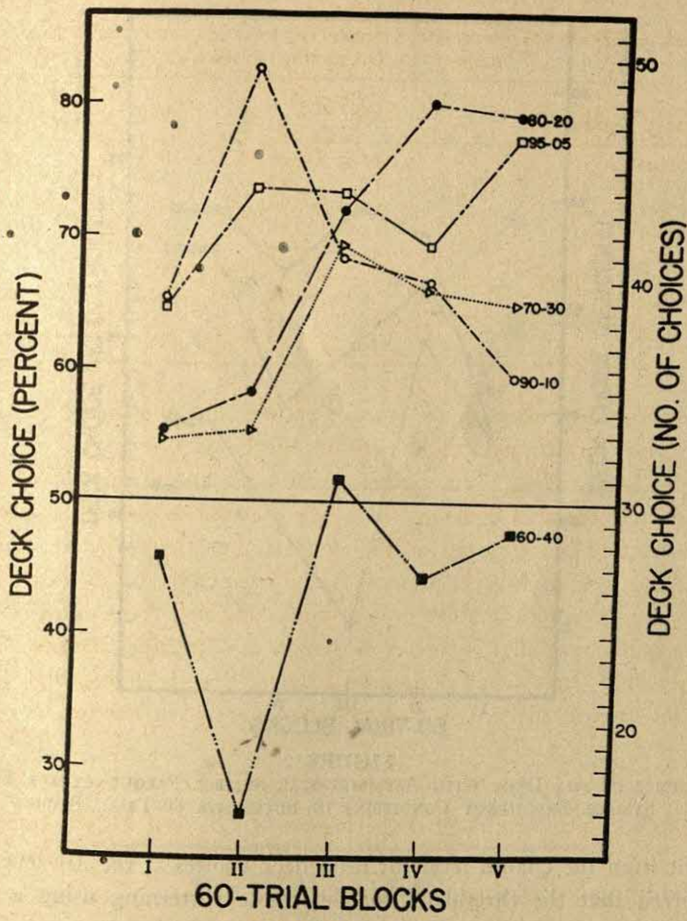


FIGURE 1
Deck Choice of the Deck with Asymmetrical Symbol Frequency for Two-Deck Symbol Frequency Conditions in Successive 60-Trial Blocks

correlation coefficients, apparently did not alter the prediction of symbols at their actual frequency of occurrence.

*T*-tests were performed on overall differences between deck choices for each condition to determine if and where a threshold of patterning occurs by use of the deck choice response measure. Deck choices for the 70-30 per cent symbol frequency condition were significantly different from chance at the .05 level for the total of the last three trial blocks ($t = 2.21$ with 17 *d.f.*). All of the more asymmetrical symbol frequencies were significantly

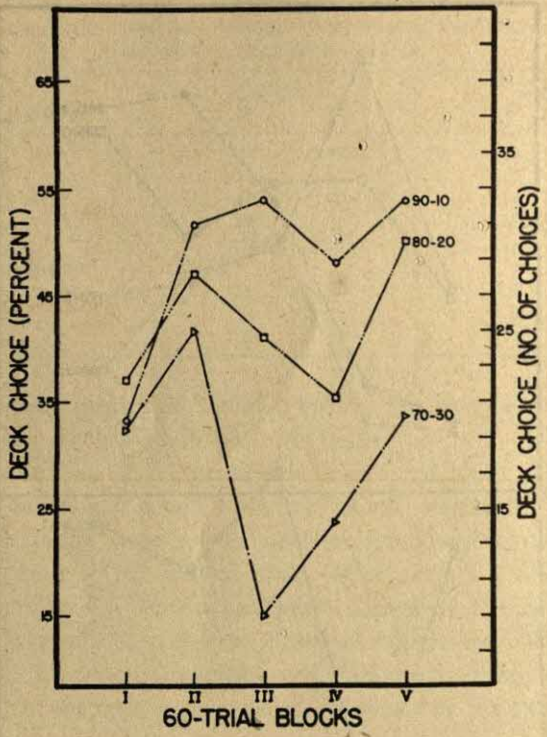


FIGURE 2
Deck Choice of the Deck with Asymmetrical Symbol Frequency for Four-Deck Symbol Frequency Conditions in Successive 60-Trial Blocks

different from the chance level for total deck choices at the .01 level. It can be inferred that the threshhold for detectable patterning using a between-series preference method in which a 50-50 per cent deck is in competition with a non-50-50 per cent deck must be in the neighborhood of a 70-30 per cent symbol frequency.

For the three symbol frequencies administered with *three* 50-50 per cent decks in Experiment *B* (Figure 2) a significant difference between conditions occurred later and more persistently than with the two-deck conditions. The two-deck conditions of Experiment *A* were significantly different only at Block II, while the four-deck conditions were significantly different for the combined totals of the last three blocks at the .01 level. ($F = 39.25$ with 2 and 51 *d.f.*)

The "balancing" conditions of Experiment *C* (Figure 3) are graphed with

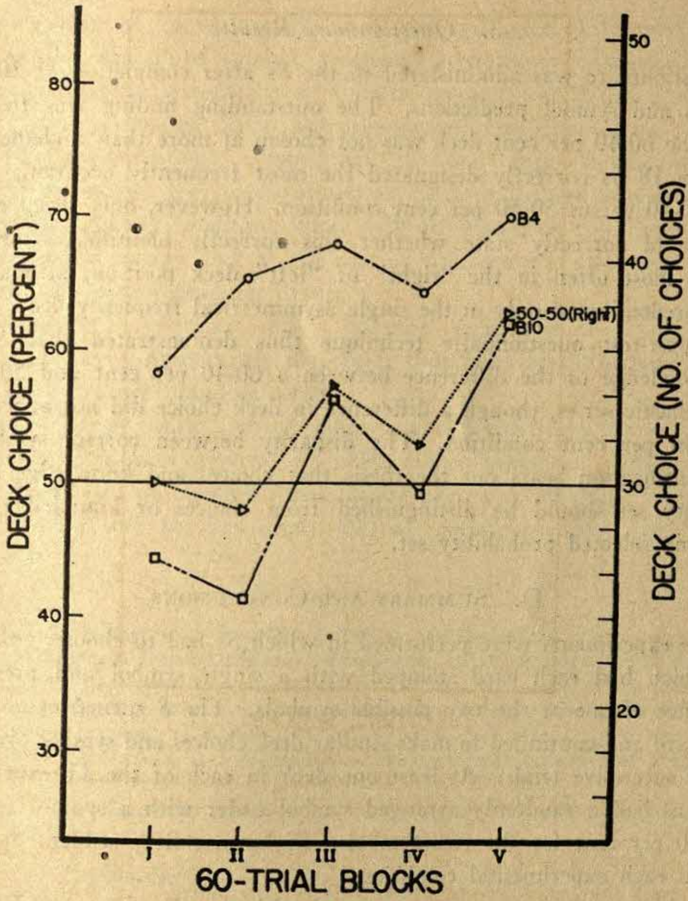


FIGURE 3

DECK CHOICE OF THE B10, B4, AND 50-50 PER CENT DECKS IN SUCCESSIVE 60-TRIAL BLOCKS

the control condition of a 50-50 per cent deck versus a 50-50 per cent deck, making the condition maximally insoluble. A slight position habit favoring the right deck develops during later choices. The selection of the "balanced-every-four" (B4) series is significantly different from chance at the .01 level ($t = 3.03$ with 17 d.f.) Apparently requiring a 50-50 per cent symbol frequency series to converge exactly every four symbols created sufficient constraint to ensure detection from a 50-50 per cent symbol frequency series converging exactly only every 100 symbols. A less constrained series (B10) converging exactly every 10 symbols was not detectable.

### 3. *Questionnaire Results*

A questionnaire was administered to the *S*s after completion of 300 deck selections and symbol predictions. The outstanding finding was that even though the 60-40 per cent deck was not chosen at more than a chance level, 14 of the 18 *S*s correctly designated the most frequently occurring symbol in the 60-40 versus 50-50 per cent condition. However, only seven of these 14 *S*s could correctly state whether this correctly identified symbol had appeared most often in the "right" or "left" deck position, although one symbol predominated only in the single asymmetrical frequency deck.

The post-test questionnaire technique thus demonstrated that *S*s have some knowledge of the difference between a 60-40 per cent and 50-50 per cent stochastic series, though a difference in deck choice did not appear until the 70-30 per cent condition. The disparity between correct symbol and deck identification bears out the thesis that choices and knowledge about a probability set should be distinguished from choices or knowledge within an already selected probability set.

## D. Summary and Conclusions

Three experiments were performed in which *S*s had to choose both a card deck which had each card stamped with a single symbol and predict the appearance of one of the two possible symbols. The *S* turned over the predicted card and continued to make similar deck choices and symbol predictions for 300 successive trials. At least one deck in each of the 11 experimental conditions had a randomly arranged symbol order with a symbol frequency of 50-50 per cent for the two symbols. Eighteen college student *S*s participated in each experimental condition.

The three experiments were concerned with the effects produced by varying proportionality, the number of competing decks, and the specification of where exact 50-50 per cent stochastic symbol frequency occurred.

Major findings were:

1. *S*s respond to long series of successive two-choice, stochastic stimuli with different symbol frequencies by predicting the symbols at the frequency of occurrence rate despite frequent sporadic and unsystematic interruptions.

2. Choice between series depends on more complex considerations than symbol prediction within a series. Although choice preference is ordered according to series proportionality up to 120 trials, subsequent trials show a decline in the preference popularity of some relatively "easy" stochastic series.

3. An estimate of approximately 70-30 per cent symbol frequency was obtained as the degree of symbol frequency asymmetry necessary for a stochasticly converging, two-choice series to be detectable in direct preference competition with a random 50-50 per cent symbol frequency series.

4. Adding two additional competitive "random" decks significantly depresses total preference for the asymmetrical deck, but there is a more systematic ordering among conditions than with the two-deck conditions. Predictions of symbols in any deck at the rate of occurrence is unaltered by the addition of two more competing decks.

5. Making 50-50 per cent series "balance" exactly over a very short interval does create a significantly detectable constraint which negates claims that such sequences are "random series."

6. Questionnaire results showed more symbol knowledge for the 60-40 per cent series than was reflected by performance measures. This within-series knowledge had no between-series counterpart since the most frequently appearing symbol was not accurately located, despite the fact there were only two series.

## REFERENCES

1. FISHER, R. A., & YATES, F. Statistical Tables for Biological, Agricultural, and Medical Research (second edition). Edinburgh: Oliver & Boyd, 1943. Pp. 112.
2. GRANT, D. A., HORNSETH, J. P., & HAKE, H. W. The influence of the inter-trial interval on the Humphreys "random reinforcement" effect during the extinction of a verbal response. *J. Exper. Psychol.*, 1950, **40**, 609-612.
3. HAKE, H. W., & HYMAN, R. Perception of the statistical structure of a random series of binary symbols. *J. Exper. Psychol.*, 1953, **45**, 64-74.
4. MOSES, L. Non-parametric statistics for psychological research. *Psychol. Bull.*, 1952, **49**, 122-143.

*Department of Psychology*
*Lawrence College*
*Appleton, Wisconsin*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 291-295.

# THE EFFECT OF NICOTINE ON MAZE BEHAVIOR OF ALBINO RATS*

*Department of Anatomy, The Chicago Medical School*

J. M. ESSENBERG

## A. INTRODUCTION

In a previous study on the effects of nicotine and cigarette smoke on pregnancy in rats (1), rather marked changes were observed in the reproductive organs as well as alterations in many other organs including the central nervous system. This latter observation prompted the present study of the effects of nicotine on the maze behavior of albino rats as tested with the Lashley 8-cul-de-sac maze.

## B. METHODS

Forty albino rats of the Wistar strain from five equal-sized litters were used for this experiment. At sexual maturity, each litter was divided into control and experimental groups with as even a representation of the sexes as possible. Both groups were then trained to a criterion of 10 consecutive errorless runs on two successive days Training was routinely carried out in the forenoon and food in the form of Purina pellets was always available in the incentive chamber at the end of the maze. Each rat was allowed to feed for about 10 seconds before it was returned to the starting chamber or removed to its home cage. Food was withheld for 20 hours before running the maze throughout the testing periods, and the weight of each animal was kept relatively constant by proper changes in the length of the feeding periods. As soon as the training criterion had been attained, the treatment of the experimental animals with nicotine began. Adopting the method used in the previous experiment (2), 0.5 cc of an 0.01 per cent solution of nicotine (pure alkaloid) in normal saline was injected intraperitoneally every day, except Sunday, into each experimental animal. Six of the control animals were injected with normal salt solution paralleling the injection of the nicotine solution in their litter mate experimental animals. The remaining control rats were left untreated.

Test runs were made from the day following the first injection. All the

---

*Received in the Editorial Office on December 11, 1953, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

rats were required to run the maze 10 times each day and time in minutes, errors, and behavior were recorded. Each time an animal entered a blind alley was scored as an error. Immediately following the beginning of the injection of the nicotine an increase in the time and error scores among the treated animals was observed. After this initial rise a decrease toward the original levels was observed which appeared to reach a fairly definite plateau somewhat above pretreatment levels. When 80 per cent of the experimental rats had attained such a plateau, all animals were given a rest period of three days with food ad lib and then maze running was renewed. Subsequent rest periods were gradually lengthened to three months where they remained until the end of the experiment. Regular injections of nicotine were continued during these rest periods. Water was always available.

The experiment was terminated at the end of two years. At that time 24 animals, 13 experimental and 11 control, remained. There had been no losses from two of the litters.

## C. Results

In spite of some fluctuations, the trend of the time and error scores of the experimental group was always toward advancing levels, while those of the control animals remained relatively constant. About a year after the experiment began, marked behavioral changes became evident in the experimental group. These animals no longer "galloped" through the maze. They began to shy from a noise and to hesitate before turning a corner as well as in front of the incentive chamber. The nicotine injected animals sometimes refused to run the maze, and would remain passively in the starting chamber or in one of the alleys or would wash their faces and other parts of their bodies. They might continue such activities for an indefinite time if undisturbed. The scores for the days the animals refused to run have been disregarded. Among the experimental animals, the number of inactive days varied but the total was never large. Such inactive days, which seldom occurred in the first year of the experiment, were confined to the nicotine injected animals and were not seen in the controls. The reason for this inactivity was impossible to determine since it did not seem to be correlated with the lack of appetite, season, etc. The total number of runs made by the individual animals was approximately the same (± 5 per cent).

The growth rates and the final weights of the experimental group were lower than those of the controls but since all of the animals had free access to food during rest periods, no apparent explanation is available. It could have been due either to an inadequate food intake or to a decrease in utilization of food. Similar observations have been made by other investigators of nicotine-treated rats (3, 5, 6, 7).

At the termination of the experiment there remained 11 control animals. Six of these were the animals that had been injected daily with 0.5 cc of normal saline solution intraperitoneally. Although it has been reported that the repeated injection of normal salt solution would damage the internal organs of the rat, particularly the gonads (4), no histological differences were found between the control animals injected with normal salt solution and those receiving no injections. Their growth rates and final weights were similar and their time and error scores in the maze were virtually the same. For this reason, the two control groups were combined in the statistical analysis of the results.

The results are summarized in Table 1. The initial scores are those of the last 20 runs of the training period which were all errorless runs and the time score is the mean of the 20 runs. The first half scores are the means of the scores of all runs during the first year of the experiment and likewise the second half scores are the means of the scores of all the runs during the second year of the experiment.

TABLE 1
MEAN SCORE OF EXPERIMENTAL AND CONTROL RATS

| | | Initial score | | First half | | Second half* | |
|---|---|---|---|---|---|---|---|
| Groups | Number | Error | Minutes | Error | Minutes | Error | Minutes |
| Experimental | 13 | 0 | 4.70 | 4.9 | 6.9 | 7.17 | 8.59 |
| Controls | 11 | 0 | 3.25 | 2.7 | 4.2 | 2.79 | 4.09 |

*Statistical analysis is given in the text.

When the time scores of the control group were compared, it was found that the differences between the initial time scores and those of either the first or second halves of the experiment were not significant. Although the control group did not continue to make errorless runs, there was no significant increase in the second half scores as compared with the first half scores. On the contrary, the first half scores and the second half scores of the experimental animals, when compared with either their own initial performances or with the performances of the control group, reveal highly significant differences. The means of the first half time scores of the experimental group are greater than their initial time scores with a confidence level of 1 per cent. The means of the second half time scores are even greater than those of the first half, and when a comparison is made between the means of the second half time scores and the initial time scores, it is found that the difference is highly significant with a confidence level of 0.1 per cent. The means of the first half time scores of the experimentals are greater than the first half time scores of the control animals and this difference, too, is significant with a confidence level of 2 per cent. Moreover the means of the second half time scores of the experimental group are considerably greater than those of the control group and this difference is highly significant with a confidence level of 0.1 per cent. Similarly the first half error scores of the experimental group are larger than those of the control group (confidence level of 1 per cent), while the greater difference between the means of the error scores for the second half of the experiment is highly significant with a confidence level of 0.1 per cent. It is thus evident that the performances of the experimental group became significantly poorer during the first half of the experiment than either their own initial performances or those of the control group and that this trend became more pronounced during the second year when all differences between initial and final scores and between experimental and control animals became significant at the 0.1 per cent level of confidence.

## D. DISCUSSION

A deterioration in maze performance of the nicotine injected animals as compared to the controls has been demonstrated but its meaning is not clear and the extent to which these findings can be extrapolated to other species is problematical. It is possible that a similar deterioration in performance might be a regular feature of any type of chronic poisoning but this remains to be determined. More difficult to assess is the generality of these results. The dosage was calculated to be the equivalent of that achieved by a human being of average weight smoking a package of cigarettes each day. The possibility that this would be sufficient to produce some deterioration in performance should at least be considered. The altered maze behavior (not mass performance in terms of time and error) is intriguing. There is apparently a decided emotional deterioration which warrants further investigation.

## E. SUMMARY

1. A weak solution of nicotine was injected intraperitoneally into a series of albino rats for a period of two years.
2. The animals were trained in an 8-cul-de-sac Lashley maze before treatment and retested at varying times throughout the experimental period.
3. The error and time scores of the experimental rats were appreciably

increased over their initial scores and over those of their litter mate controls during the first year of the experiment. These differences were shown to be significant by the "*t*" test. This trend became more marked in the final year when all differences btween initial and final scores of the experimental and control animals became significant at the 0.1 per cent level of confidence.

4. Marked behavioral changes occurred in the experimental group. These did not develop until early in the second year of the experiment.

5. Growth rate in the experimental groups was lower than in the controls.

## REFERENCES

1. ESSENBERG, J. M., SCHWIND, JUSTIN V., & PATRAS, ANNE R. The effect of nicotine and cigarette smoke on pregnant female albino rats and their offspring. *J. Lab. Clin. Med.*, 1940, **25**, 708-717.
2. ESSENBERG, J. M., FAGAN, L., & MALERSTEIN, A. J. Chronic poisoning of the ovaries and testes of albino rats and mice by nicotine and cigarette smoke. *West. J. Surg., Obst., Gynec.*, 1951, **59**, 27-32.
3. FIELD, H.E. The immediate effects of tobacco smoke on the activity of rats. *Univ. Calif. Publ. Physiol.*, 1926, **5**, 189-194.
4. HOTOVY, R. Versuche zur Frage des Einflusses der chronischen Nikotinschädigung auf Fruchtbarkeit und Nachkommenschaft. *Arch. Exper. Path. u. Pharmak.*, 1948, **205**, 54-56.
5. NAKASAWA, R. Der Einfluss der chronischen Nikotinvergiftung auf die Funktion der Geschlechtsorgane der weiblichen Ratten. *Jap. J. Med., S.C.*, 1933, **7**, 1-37.
6. PECKSTEIN, I. A., & REYNOLDS, W. R. The effect of tobacco smoke on the growth and learning behavior of the albino rat and its progeny. *J. Comp. Psychol.*, 1937, **24**, 459-469.
7. PHILLIPS, H. C. The immediate effects of tobacco smoke on the learning ability of albino rats. *J. Comp. Psychol.*, 1937, **24**, 471-486.

*Department of Anatomy*
*Chicago Medical School*
*710 South Wolcott Avenue*
*Chicago 12, Illinois*

Published as a separate and in *The Journal of Psychology,* 1954, **37**, 297-305.

# SCALE POSITION AND PERFORMED MELODIC OCTAVES*[1]

*Department of Psychology, The Pennsylvania State University*

JOHN F. CORSO[2]

## A. INTRODUCTION

A considerable amount of experimental research has been directed toward the investigation of the psychological dimensions of tonal experience (10). For the most part, however, this research has been concerned with the dimensions of *pure* tones. Little information is currently available on the subjective attributes or psychophysical relationships of *complex* musical sounds. Until further investigations are undertaken, it remains unknown whether the results obtained for pure tones can be legitimately extended to cover the typical sounds of music. One of the fundamental problems in this area is the determination of a pitch function for complex tones and the establishment of subjective magnitudes for conventional musical intervals.[3]

The data of Stevens, Volkmann, and Newman (17) provide some information along these lines for the case of pure tones. From the results obtained by testing five observers at 10 different frequencies by the method of fractionation, a scale was developed showing the relationship between pitch and frequency. This pitch (mel) scale was then used to measure the subjective magnitude of certain musical intervals. Specifically, it was found that for pure tones musical intervals become subjectively larger as frequency increases, up to approximately 4000 *cps;* beyond this, perceived magnitude decreases with increasing frequency. It was determined that the subjective size of different octaves[4] may vary as much as 20-to-1.

Stevens and Davis (15, p. 84) indicate that these findings support Stumpf's (18, p. 250) position that upper octaves are perceptually larger than lower

---

*Received in the Editorial Office on January 9, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Paper presented at the Twenty-fourth Annual Meeting of the Eastern Psychological Association held at Boston, Massachusetts, on April 24-25, 1953.

[2] The data of this study were collected in the Psychology Laboratory of the State University of Iowa, Iowa City, Iowa.

[3] An interval is defined as the difference in pitch between any two tones which are sounded together (harmonic interval) or successively (melodic interval) and is usually expressed as the ratio of the two physical frequencies producing the tones (19, p. 143).

[4] An octave is the interval between two frequencies having a ratio of two-to-one (12, p. 5).

octaves. Further, the data contradict the notion that "equal ratios of frequency give rise to equal intervals of pitch" as stated by Davis (1, p. 235). Pratt (9) has also shown that for pure tones a given interval increases in apparent size with an increase in the frequency of the interval tones.

Railsback (11), in a study on the tuning of pianos, has obtained data which indicate that a complex relationship might well exist between the pitch and frequency of *musical* sounds. The results of six expert tuners on a total of 16 grand and 12 upright pianos clearly showed that all tuners "stretched"[5] the octaves in both the lower and upper registers of the piano keyboard. Lewis (4) has interpreted these results as pointing toward a non-linear relation between the pitch of complex tones and the logarithm of frequency.

In view of these findings and the paucity of experimental data on psychophysical functions of complex tones, the question arises as to whether instrumental musicians, when asked to perform melodic octaves over the equal tempered scale, exhibit a tendency similar to that of piano tuners in which deviations from the established octave requirement of a two-to-one frequency ratio are observed. If, as some investigators (2, 13) have concluded, the pitch of a complex tone is determined by the frequency of its fundamental, significant changes in physical octave extent are to be expected with changes in the fundamental frequencies of the octave tones. This follows from existing pure tone data (14, 16) which demonstrate the existence of a non-linear relation between pitch and frequency. Octave intervals performed over the range of the equal tempered scale are expected, therefore, to decrease in physical magnitude with increases in the frequency of the octave tones, provided the physical octaves are adjusted on the basis of subjective octaves of equal magnitude.

Specifically, the present experiment was undertaken to determine the manner in which melodic octaves performed on musical instruments varied in physical extent with shifts in the location of the octave tones along the frequency continuum of the equal tempered scale.

## B. METHODOLOGY

### 1. *Apparatus*

The apparatus of the study consisted of two units: (*a*) a Conn chromatic stroboscope used by the experimenter (*E*) to measure the frequency of each note played by the subject ($S$), and ($b$) a footswitch-activated signal light used by $S$ to denote to $E$ the proper moment at which the frequency reading should be taken. All measurements were made in a sound deadened chamber and no attempt was made to control the intensity of the notes played.[6]

---

[5] A "stretched" octave may be defined as an octave having more than 1200 cents, where one cent represents 1/100 of an equal tempered semitone. Since there are 100 cents per semitone and 12 semitones per octave, an equal tempered semitone must contain 1200 cents (12, pp. 26-27).

### 2. *Procedure*

The experimental plan of the study is shown in summarized form in Table 1. There were four major trial series, two in which the melodic octaves were performed in ascending chromatic order and two in which the octaves were performed in descending chromatic order. For each of these orders, two sequences were followed in playing the successive notes which formed the octaves. Half the time the lower note of the interval was played first; the other half the time, the upper note was played first. Trial Series 1 and 2 were run at the first test hour; trial Series 3 and 4, at the second test hour. In an attempt to obtain more reliable data, the ascending-descending sequence of chromatic scale performance was retained at both test hours to provide $S$s with a "practice" or "warm-up" period before playing the higher and more difficult octaves.

TABLE 1
EXPERIMENTAL PLAN

| Test hour | Trial series | Direction of chromatic scale | Sequence of notes played in forming melodic octaves |
|---|---|---|---|
| 1 | 1 | Ascending | Lower — Upper |
|  | 2 | Descending | Upper — Lower |
| 2 | 3 | Ascending | Upper — Lower |
|  | 4 | Descending | Lower — Upper |

The $S$s of this study were five highly trained musicians, all of whom were members of the State University of Iowa symphony orchestra. These included one bassoonist, one alto saxophonist, one violinist, and two clarinetists. Each $S$ served for a single one-hour session on each of two successive days. Individual measurements were made for each $S$ over that portion of the instrumental register in which extreme confidence was displayed of accurate intonation.

---

[6] The data available indicate that intensity has little effect on the pitch of complex tones (3, 5).

## C. RESULTS

In the treatment of data, each *S* was assigned a performance score for each of the four trials at each position[7] of the equal tempered scale at which octave measurements were made. The assigned score, expressed in cents, was the difference between the performed octave and the equal tempered octave of 1200 cents. For bassoon, saxophone, and violin these "error" scores were in each case arranged in a trials by scale positions analysis of variance table and the total sum of squares was analyzed into two components. For clarinets, mean "error" scores for four trials were computed for each *S* at each scale position tested and the data were arranged in a subjects by scale positions analysis of variance table. In this case, the total sum of squares was analyzed into three components.

TABLE 2
SUMMARIES OF ANALYSIS OF VARIANCE FOR TESTING TREND IN MEANS

| Instrument | Source of variance | Sum of squares | *df* | Mean square | *F* ratio |
|---|---|---|---|---|---|
| Bassoon | B/scale positions | 26709.80 | 28 | 953.92 | 9.19* |
| | W/scale positions | 9023.50 | 87 | 103.72 | |
| | Total | 35733.30 | 115 | | |
| Alto-Saxophone | B/scale positions | 3629.80 | 17 | 213.52 | 3.92* |
| | W/scale positions | 2940.75 | 54 | 54.46 | |
| | Total | 6570.55 | 71 | | |
| Violin | B/scale positions | 4932.50 | 24 | 205.52 | 2.02** |
| | W/scale positions | 7618.25 | 75 | 101.57 | |
| | Total | 12550.75 | 99 | | |
| Clarinet | B/scale positions | 7276.58 | 24 | 303.19 | 3.84* |
| | B/subjects | 31.20 | 1 | 31.20 | |
| | Residual | 1892.23 | 24 | 78.84 | |
| | Total | 9200.01 | 49 | | |

*$p < .01$.
**$p < .02$.

Table 2 presents a summary of the data obtained for each analysis performed to test the null hypothesis of trend in means for the several instruments studied. Specifically stated, the hypothesis tested in each case was that, for the population involved, changes in scale position are not accompanied by changes in physical octave extent, i.e., the regression line of population means is a horizontal straight line. The test for trend in means based on a single sample followed the procedure outlined by Lindquist (6).

---

[7] The term "position" as used in this study refers to the equal tempered scale location of the two tones defining an octave and could be specified by a pair of letters expressed in conventional musical notation where Middle C corresponds to $C_4$. For simplicity, however, scale position is herein expressed by a *single number* equal to the geometric mean of the fundamental frequencies of the two tones forming the octave; e.g., the scale position for the octave formed by $C_3$ and $C_4 = \sqrt{(130.8)\ (261.6)} = 186$ *cps.*

As shown in Table 2, the obtained *F* ratios and their levels of significance are as follows: (*a*) for bassoon, the *F* ratio was 9.19 which, for 28 and 87 degrees of freedom (*df*), is significant at the one per cent level; (*b*) for alto saxophone, the *F* ratio was 3.92 which, for 17 and 54 *df*, is significant at the one per cent level; (*c*) for violin, the *F* ratio was 2.02 which, for 24 and 75 *df*, is significant at the two per cent level; and (*d*) for clarinets, the *F* ratio was 3.84 which, for 24 and 24 *df*, is significant at the one per cent level. The null hypothesis of trend in means was rejected in each case.

Figure 1 presents in graphical form the empirical function obtained for each of the several instruments tested. Scale position, as previously defined, is plotted in *cps* on the logarithmic abscissa; mean octave "error" in cents above and below equal temperament is plotted on the ordinate. Part *A* of Figure 1 shows the function obtained for bassoon; Part *B*, for alto saxophone; Part *C*, for violin; and Part *D*, for the two clarinets. The horizontal zero line in each plot represents equal temperament, or octaves of 1200 cents. If the octaves measured had been performed in accordance with the requirements of equal temperament, all means would have fallen on this line, except for chance deviations.

Since the analysis of variance results indicated that the null hypothesis of trend in means was untenable in each case tested and that the discrepancies from equal temperament noted in Figure 1 were not attributable to random fluctuations, a second hypothesis was tested that a linear relationship obtains between scale position and physical octave extent, i.e., increases in the independent variable are accompanied by proportional changes in the dependent variable. For each instrument, a test for departure from linearity was performed using the analysis of covariance techniques described by Lindquist (7, pp. 235-238). For calculating the sums of squares required in the analysis, scale positions were arbitrarily assigned numerical values in arithmetic sequence, with the number one assigned in each case to the lowest melodic octave.

A summary of the results obtained from the analysis of covariance for each instrument is presented in Table 3. The obtained *F* ratios and their significance levels are as follows: (*a*) for bassoon, the *F* ratio was 7.94 which, for 27 and 87 *df*, is significant at the one per cent level; (*b*) for alto saxophone, the *F* ratio was 4.16 which, for 16 and 54 *df*, is significant at the one per cent level; (*c*) for violin, the *F* ratio was 2.07 which, for 23 and 75 *df*,

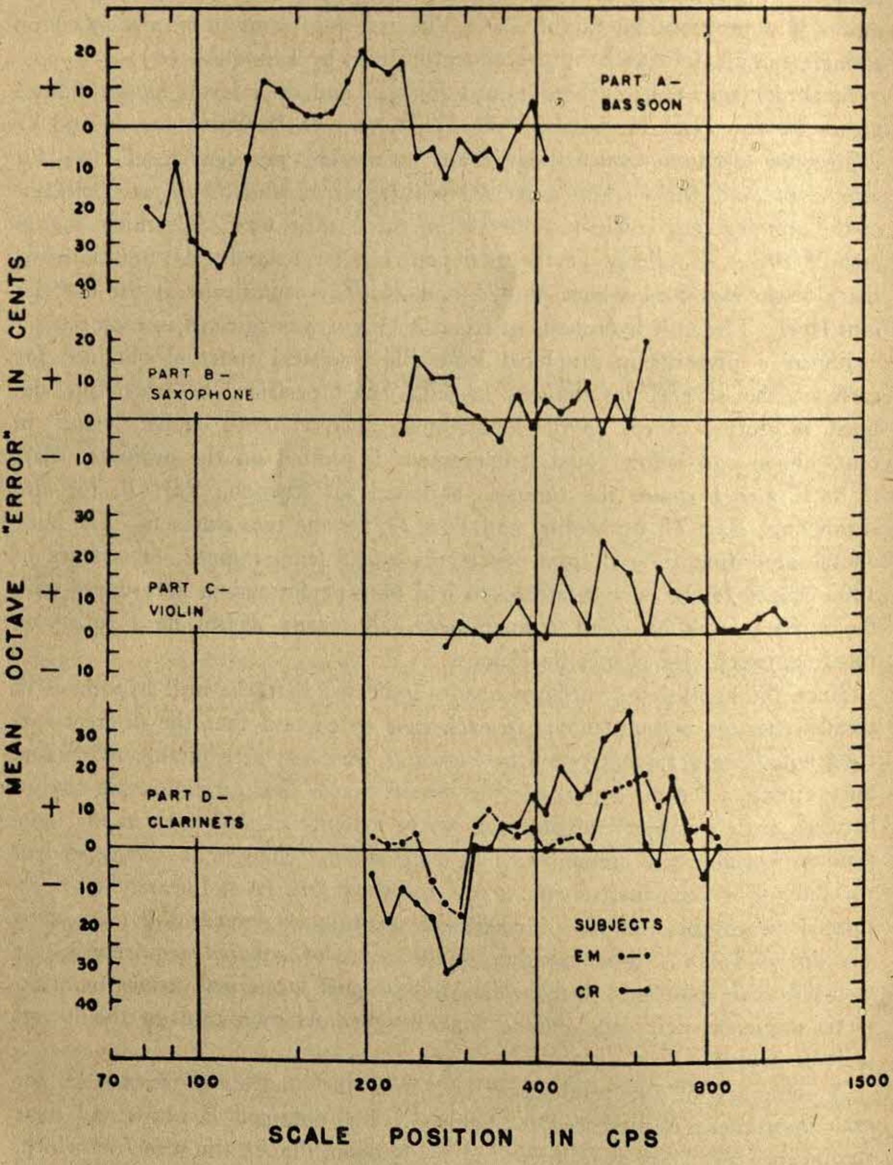


FIGURE 1

Differences between the Physical Extent of Performed Melodic Octaves and Equal Tempered Octaves of 1200 Cents: Each Point on Each Curve Represents the Mean of Four Trials at Each Designated Scale Position

TABLE 3
SUMMARIES OF ANALYSIS OF COVARIANCE FOR TESTING LINEARITY OF REGRESSION

| Instrument | Source of variance | Sum of squares | *df* | Mean square | *F* ratio |
|---|---|---|---|---|---|
| Bassoon | Linear regression | 4461.42 | 1 | | |
| | Departure from linearity | 22247.38 | 27 | 823.98 | 7.94* |
| | W/scale positions | 9023.50 | 87 | 103.72 | |
| | Total | 35733.30 | 115 | | |
| Alto-Saxophone | Linear regression | 8.52 | 1 | | |
| | Departure from linearity | 3621.28 | 16 | 226.33 | 4.16* |
| | W/scale positions | 2940.75 | 54 | 54.46 | |
| | Total | 6570.55 | 71 | | |
| Violin | Linear regression | 98.86 | 1 | | |
| | Departure from linearity | 4833.64 | 23 | 210.16 | 2.07* |
| | W/scale positions | 7618.25 | 75 | 101.57 | |
| | Total | 12550.75 | 99 | | |
| Clarinets | Linear regression | 2627.07 | 1 | | |
| | Departure from linearity | 4649.51 | 23 | 202.15 | 2.56** |
| | B/subjects | 31.20 | 1 | | |
| | Residual | 1892.23 | 24 | 78.84 | |
| | Total | 9200.01 | 49 | | |

*$p < .01$.
**$p < .02$.

is significant at the one per cent level; and (*d*) for clarinets, the *F* ratio was 2.56 which, for 23 and 24 *df*, is significant at the two per cent level. The results indicate that, for the populations tested, increases in the scale position of performed melodic octaves are not accompanied by proportional changes in physical, octave extent, i.e., the hypothesis of linearity is untenable in each case.

## D. DISCUSSION

It appears, therefore, that a significant tendency exists among the musicians of this study to alter the physical extent of performed melodic octaves with changes in the location of the octaves along the frequency continuum of the equal tempered scale. Further, the relationship which obtains between scale position and octave extent cannot be expressed by a simple linear function. In general, from Figure 1, it may be stated that the magnitude of physical octaves tends to increase with upward shifts in the placement of the octaves on the frequency continuum. With the exception of the alto saxophone, maximal octave "stretching" occurs in the midrange portion of each instrumental register studied. This effect, however, is probably related to the physical limitations of tone production characteristic of most musical instruments.

These findings, while indicating the existence of performed melodic octaves of *unequal* magnitude, do not support the hypothesis that performed octaves become (physically) *smaller* with increases in frequency throughout the musical range. Apparently, either (*a*) the pure tone pitch function does not adequately represent the pitch function for complex musical sounds, or (*b*) if it does, physical octaves for complex sounds are not adjusted on the basis of subjective octaves of equal magnitude. Although a full explanation of the results of this study cannot be given at this time, the data suggest a non-linear relation between the pitch of complex tones and the logarithm of frequency. It is emphasized that while the relationship between pitch and frequency is probably the most important factor in the explanation of the present results, the pitch of musical sounds may also be dependent upon other physical factors, such as harmonic structure, duration, and vibrato (8, pp. 246-248).

## E. SUMMARY

In an attempt to determine the manner in which melodic octaves performed on musical instruments varied in physical extent with shifts in the location of the octaves along the frequency continuum, data were obtained for five highly trained subjects (a violinist, a bassoonist, an alto saxophonist, and two clarinetists) over a wide portion of the equal tempered scale. Treatment of the data involved tests for trends in mean octave extents and tests for linearity of means through appropriate applications of the methods of analysis of variance and covariance, respectively. For each instrument studied, the results indicated a significant trend in mean octave extent to increase with shifts in the location of the performed octaves along the frequency continuum. The hypothesis of linearity of mean octave extent and scale position, however, was in no case tenable. The interpretation of results suggests that the pitch of musical sounds is not directly proportional to the logarithm of frequency and is probably complexly conditioned.

## REFERENCES

1. DAVIS, A. H. Modern Acoustics. New York: Macmillan, 1934. Pp. 345.
2. FLETCHER, H. Newer concepts of the pitch, the loudness, and the timbre of musical tones. *J. Franklin Inst.,* 1935, **220**, 405-429.
3. ————. Loudness, pitch, and timbre of musical tones and their relation to the intensity, the frequency, and the overtone structure. *J. Acoust. Soc. Amer.,* 1934, **6**, 59-69.
4. LEWIS, D. Pitch as a psychological phenomenon. *Proc. Music Teachers Nat. Assoc.,* 1939, 121-133.
5. LEWIS, D., & COWAN, M. The influence of intensity on the pitch of violin and 'cello tones. *J. Acoust. Soc. Amer.,* 1936, **8**, 20-22.

6. LINDQUIST, E. F. Goodness of fit of trend curves and significance of trend differences. *Psychometrika*, 1947, **12**, 65-78.

7. ———. Statistical Analysis in Educational Research. New York: Houghton Mifflin, 1940. Pp. 266.

8. OLSON, H. F. Musical Engineering. New York: McGraw-Hill, 1952. Pp. 369.

9. PRATT, C. C. Comparison of tonal distances. *J. Exper. Psychol.*, 1928, **11**, 77-87.

10. PSYCHOACOUSTIC LABORATORY, Harvard University. A Bibliography in Audition. Cambridge: Harvard Univ. Press, 1950.

11. RAILSBACK, O. L. A study of the tuning of pianos. *J. Acoust. Soc. Amer.*, 1938, **10**, 86.

12. SCHOEN, M. The Psychology of Music. New York: Ronald Press, 1940. Pp. 258.

13. SCHOUTEN, J. F. The perception of pitch. *Philips Tech. Rev.*, 1940, **5**, 286-294.

14. SHOWER, E. G., & BIDDULPH, R. Differential pitch sensitivity of the ear. *J. Acoust. Soc. Amer.*, 1931, **3**, 275-287.

15. STEVENS, S. S., & DAVIS, H. Hearing: Its Psychology and Physiology. New York: Wiley, 1938. Pp. 489.

16. STEVENS, S. S., & VOLKMANN, J. The relation of pitch to frequency: a revised scale. *Amer. J. Psychol.*, 1940, **53**, 329-353.

17. STEVENS, S. S., VOLKMANN, J., & NEWMAN, E. B. A scale for the measurement of the psychological magnitude pitch. *J. Acoust. Soc. Amer.*, 1937, **8**, 185-190.

18. STUMPF, C. Tonpsychologie. Leipzig: Hirzel, 1883 (Vol. 1). Pp. 427.

19. WARREN, H. C. Dictionary of Psychology. New York: Houghton Mifflin, 1934 Pp. 371.

*Department of Psychology*
*Pennsylvania State University*
*State College, Pennsylvania*

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

## The General Field of Psychology

EDITED BY

CARL MURCHISON

Gordon Allport
Harvard University

Arthur G. Bills
University of Cincinnati

E. A. Bott
University of Toronto

Leonard Carmichael
Smithsonian Institution

Herbert S. Conrad
U. S. Office of Education

Elmer K. Culler
University of Rochester

John F. Dashiell
University of North Carolina

Roy M. Dorcus
University of California at Los Angeles

Horace B. English
Ohio State University

Franklin Fearing
University of California at Los Angeles

Frank A. Geldard
University of Virginia

Clarence H. Graham
Columbia University

Ernest R. Hilgard
Stanford University

H. M. Johnson
Tulane University

Heinrich Klüver
University of Chicago

Carney Landis
Columbia University

K. S. Lashley
Harvard University

Norman L. Munn
Bowdoin College

John Paul Nafe
Washington University

Donald G. Paterson
University of Minnesota

Lee Edward Travis
University of Southern California

Robert C. Tryon
University of California

Michael J. Zigler
Wellesley College

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

## VOLUME 38

1954

Copyright, 1954, by The Journal Press
Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

# AUTHOR INDEX

Anderson, Scarvia B. ... 151
Aumack, Lewis ... 79
Bartley, S. Howard ... 313, 453
Basowitz, Harold ... 495, 503
Beaver, Alma P. ... 73
Bendig, A. W. ... 421
Berg, Irwin A. ... 475
Bjerstedt, Ake ... 67
Boder, David P. ... 3
Brodbeck, Arthur J. ... 271
Catalano, Frank L. ... 203
Clausen, Johs ... 211
Comrey, Andrew L. ... 251
Craig, Eugene A. ... 265
Deane, G. ... 235
Ekman, Gasta ... 467
Fonda, Charles P. ... 427
Gilmore, James L. ... 359
Gogel, Walter C. ... 483
Grinker, Roy R. ... 503
Grossack, Martin M. ... 127
Hastings, Philip H. ... 301
Hess, Eckhard H. ... 483
Hochberg, Julian E. ... 83
Hughes, J. B., II ... 421
Hymovitch, Bernard ... 281
Jackson, Douglas N. ... 339
Karson, Samuel ... 321
Katcher, Allan ... 89
King, Gerald F. ... 383
Korchin, Sheldon J. ... 495, 503
Krueger, Erick G. ... 191
Landis, Carney ... 211
Lane, John E. ... 457
Lawrence, Douglas H. ... 367
Levinson, Daniel J. ... 103
Luchins, Abraham S. ... 439
McCarthy, Dorothea ... 203
Mech, Edmund ... 71
Miller, James W. ... 453
Moore, Omar Khayyam ... 151
Parsons, Oscar A. ... 457
Perlmutter, Howard V. ... 131, 223, 271, 281, 291
Phillips, Leslie ... 457
Price, Pliny A. ... 191
Rapaport, Gerald M. ... 475
Riopelle, Arthur J. ... 331
Rosenman, Stanley ... 389
Schmidt, Herman O. ... 427
Seibel, Jean L. ... 313
Shaklee, A. B. ... 97
Shaw, Marvin E. ... 139
Smith, Donald E. P. ... 83
Sprague, J. L. ... 421
Stern, Ferdinand ... 321
Teuber, Hans-Lukas ... 161, 177, 191
Turner, Merle B. ... 265
Wapner, Seymour ... 51
Watson, George ... 251
Wegner, N. ... 235
Weinstein, Sidney ... 161, 177
Werner, Heinz ... 51
Zeaman, D. ... 235

# TABLE OF CONTENTS

The impact of catastrophe: I. Assessment and evaluation . . . . . . 3
By David P. Boder

Studies in physiognomic perception: I. Effect of configurational dynamics and meaning-induced sets on the position of the apparent median plane 51
By Heinz Werner and Seymour Wapner

Testproximal analysis, terminology absorption, and conceptual purification. 67
By Ake Bjerstedt

A brief note on learning "principles" and reading . . . . . . . 71
By Edmund Mech

Dominance in the personality of the student nurse as measured by the *A-S* Reaction Study . . . . . . . . . . . . 73
By Alma P. Beaver

Misconceptions concerning the interpretation of sub-group variations within normative data . . . . . . . . . . . . 79
By Lewis Aumack

The effect of "punishment" (electric shock) on figure-ground perception . . 83
By Donald E. P. Smith and Julian E. Hochberg

Some anthropomorphic comments on latent learning . . . . . . . 89
By Allan Katcher

One- and two-sided statistical tests and objectives of experimentation . . 97
By A. B. Shaklee

The intergroup relations workshop: Its psychological aims and effects . . 103
By Daniel J. Levinson

Perceived Negro group belongingness and social rejection . . . . 127
By Martin M. Grossack

Relations between the self-image, the image of the foreigner, and the desire to live abroad . . . . . . . . . . . . . . 131
By Howard V. Perlmutter

Group structure and the behavior of individuals in small groups . . . 139
By Marvin E. Shaw

Modern logic and tasks for experiments on problem solving behavior . . 151
By Omar Khayyam Moore and Scarvia B. Anderson

Judgment of visual and postural vertical after brain injury . . . . 161
By Hans-Lukas Teuber and Sidney Weinstein

Performance on a formboard-task after penetrating brain injury . . . 177
By Hans-Lukas Teuber and Sidney Weinstein

Tactile extinction in parietal lobe neoplasm . . . . . . . . 191
By Erick G. Krueger, Pliny A. Price, and Hans-Lukas Teuber

Infant speech as a possible predictor of later intelligence . . . . . 203
By Frank L. Catalano and Dorothea McCarthy

Certain effects of mescaline and lysergic acid on psychological functions . . 211
By Carney Landis and Johs Clausen

Impressions of influential members of discussion groups . . . . . 223
By Howard V. Perlmutter

Amplitude and latency characteristics of the conditioned heart response . . 235
By D. Zeaman, G. Deane, and N. Wegner

Nutritional replacement for mental illness . . . . . . . . . 251
By George Watson and Andrew L. Comrey

The effect of figural reproduction on recognition . . . . . . . 265
By Merle B. Turner and Eugene A. Craig

Self-dislike as a determinant of marked ingroup-outgroup preferences . . 271
By Arthur J. Brodbeck and Howard V. Perlmutter

Communication to high and low power foreign and domestic authorities . 281
By Howard V. Perlmutter and Bernard Hymovitch

Some characteristics of the xenophilic personality . . . . . . . 291
By Howard V. Perlmutter

The non-voter in 1952: A study of Pittsfield, Massachusetts . . . . 301
By Philip K. Hastings

A further study of entoptic stray light . . . . . . . . 313
By S. Howard Bartley and Jean L. Seibel

Motivation and attention: A methodological problem* . . . . . 321
By Ferdinand Stern and Samuel Karson

Facilities of the Emory University Primate Behavior Laboratory . . . 331
By Arthur J. Riopelle

A further examination of the rôle of autism in a visual figure-ground relationship . . . . . . . . . . . . . . 339
By Douglas N. Jackson

Recall of success and failure as a function of subject's threat interpretations 359
By James L. Gilmore

The evaluation of training and transfer programs in terms of efficiency measures . . . . . . . . . . . . . . 367
By Douglas H. Lawrence

Research with neuropsychiatric samples . . . . . . . . . 383
By Gerald F. King

The multiplicity of the normal ego structure . . . . . . . . 389
By Stanley Rosenman

Anxiety level, response alternation, and performance in serial learning . . 421
By J. B. Hughes, II, J. L. Sprague, and A. W. Bendig

Rorschach scores in the manic state . . . . . . . . . . 427
By Hermann O. Schmidt and Charles P. Fonda

The relation of size of light to autokinetic effect . . . . . . . 439
By Abraham S. Luchins

Some circumstances surrounding apparent movement in the line of regard . 453
By S. Howard Bartley and James W. Miller

Performance on the same psychomotor task under different stressful conditions 457
By Oscar A. Parsons, Leslie Phillips, and John E. Lane

Dimensions of color vision . . . . . . . . . . . . 467
By Gosta Ekman

Response bias in an unstructured questionnaire . . . . . . . . 475
By Irwin A. Berg and Gerald M. Rapaport

Natural preferences of the chick for objects of different colors . . . 483
By Eckhard H. Hess and Walter C. Gogel

Perceptual adequacy in a life stress . . . . . . . . . . 495
By Sheldon J. Korchin and Harold Basowitz

Anxiety in a life stress . . . . . . . . . . . . . 503
By Harold Basowitz, Sheldon J. Korchin, and Roy R. Grinker

---

*On p. 328, fourth line of Summary, the word "reverse" should be "same."

$20.00 per annum
$10.00 per volume
$5.00 per half volume

QUARTERLY
Two volumes per year
Immediate publication

July, 1954
Volume 38, First Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

## The General Field of Psychology

---

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

## JULY, 1954

**(Manuscripts are printed in the order of final acceptance)**

The impact of catastrophe: I. Assessment and evaluation . . . . . 3
By David P. Boder

Studies in physiognomic perception: I. Effect of configurational dynamics and meaning-induced sets on the position of the apparent median plane 51
By Heinz Werner and Seymour Wapner

Testproximal analysis, terminology absorption, and conceptual purification. 67
By Ake Bjerstedt

A brief note on learning "principles" and reading . . . . . . 71
By Edmund Mech

Dominance in the personality of the student nurse as measured by the *A-S* Reaction Study . . . . . . . . . . . . . 73
By Alma P. Beaver

Misconceptions concerning the interpretation of sub-group variations within normative data . . . . . . . . . . . . 79
By Lewis Aumack

**(OVER)**

---

Copyright, 1954, by The Journal Press
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter January 26, 1937, at the post-office at
Provincetown, Mass., under the Act of March 3, 1879

The effect of "punishment" (electric shock) on figure-ground perception . . 83
By Donald E. P. Smith and Julian E. Hochberg

Some anthropomorphic comments on latent learning . . . . . . . 89
By Allan Katcher

One- and two-sided statistical tests and objectives of experimentation . . 97
By A. B. Shaklee

The intergroup relations workshop: Its psychological aims and effects . . 103
By Daniel J. Levinson

Perceived Negro group belongingness and social rejection . . . . 127
By Martin M. Grossack

Relations between the self-image, the image of the foreigner, and the desire to live abroad . . . . . . . . . . . . . . 131
By Howard V. Perlmutter

Group structure and the behavior of individuals in small groups . . . 139
By Marvin E. Shaw

Modern logic and tasks for experiments on problem solving behavior . . 151
By Omar Khayyam Moore and Scarvia B. Anderson

Judgment of visual and postural vertical after brain injury . . . . 161
By Hans-Lukas Teuber and Sidney Weinstein

Performance on a formboard-task after penetrating brain injury . . . 177
By Hans-Lukas Teuber and Sidney Weinstein

Tactile extinction in parietal lobe neoplasm . . . . . . . . 191
By Erick G. Krueger, Pliny A. Price, and Hans-Lukas Teuber

Infant speech as a possible predictor of later intelligence . . . . . 203
By Frank L. Catalano and Dorothea McCarthy

Certain effects of mescaline and lysergic acid on psychological functions . . 211
By Carney Landis and Johs Clausen

Impressions of influential members of discussion groups . . . . . 223
By Howard V. Perlmutter

Amplitude and latency characteristics of the conditioned heart response . . 235
By D. Zeaman, G. Deane, and N. Wegner

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 3-50.

# THE IMPACT OF CATASTROPHE: I. ASSESSMENT AND EVALUATION*[1]

*Department of Psychology, University of California*

DAVID P. BODER[2]

## A. INTRODUCTION

The present article is the first in a series of four, devoted to the analysis of magnetically recorded interviews with disaster sufferers. It deals with the content analyses and traumatic evaluation of wire recorded interviews from a collection obtained in displaced persons camps and shelter houses of Europe during the summer of 1946, i.e., at about the beginning of the second year following VE-day. The *second* article is devoted to the linguistic aspects of the narratives. The *third* deals with theoretical and generalized methodological considerations, and the *fourth*—a report on a separate study—is to be devoted to the analysis of magnetically recorded interviews obtained from sufferers of the Kansas City flood of 1951 within seven to 10 days after the onset of the disaster. The latter study also presents some broad bases for comparison of traumatizations inflicted by a *nature-made* disaster in a friendly and intact environment vs. *man-made* catastrophe in a hostile environment. It seems reasonable to hypothesize that somewhere between these two reference points are located (*a*) *man-made* disasters within a *friendly* and reasonably *stable* environment such as the blitz on Coventry, and (*b*) *man-made* disasters capable of *catastrophic disorganization* of the

---

*Received in the Editorial Office on February 2, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported (in part) by a research grant (M-156 and M-582) from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

[2] *Acknowledgments:* Thanks are due for various aspects of coöperation to the following persons: Alice Brown (11), Polly Hammond (17), Audrey Uher (33), all graduate students of Illinois Institute of Technology, who were most helpful in the preliminary phases of the project while writing their Masters dissertations on topics pertaining to this study; to Frank Smith, who devoted a semester of undergraduate research at I.I.T. to the same topic; to Dr. Dora N. Boder, who constructed a most comprehensive index of a series of DP interviews, including the 10 used in this article; to LeRoy Hackworth, Donald Procter, and Richard Timmer, graduate students of psychology at U.C.L.A. who participated in the content analysis; to Norma Segal and Eleanor A. Smith who coöperated in a painstaking fashion with the typing of the manuscript. Procter also assisted with the final organization of the material. Thanks are also due for many helpful courtesies to Drs. Phil and Harriett Shurrager of I.I.T., and to Dr. Joseph Gengerelli and the Department of Psychology of the University of California at Los Angeles.

environment, such as may possibly be foreseen in the case of a thermonuclear attack.

## B. Design and Methods

The present article describes the design and tentative use of two tools—the Traumatic Inventory, and the Scale of Traumatic Values. By means of these two instruments, the contents of the interviews were analyzed as to the presence and frequency of specific traumatic experiences, and subsequently evaluated as to the intensity of the traumatic impact. The application of these devices renders six raw scores or Indices for each interview. These are:

1. *PTF—personal traumatic frequency.* This score covers the instances in which the narrator personally was the target of the trauma, alone or in conjunction with others.
2. *MTF—milieu traumatic frequency*—the traumata in which "other" specified persons or groups were the target.
3. *TLF—traumatic load frequency,* covering the whole traumatic content of the interview as expressed by the sum of *PTF* plus *MTF.*
4. *PTV—personal traumatic value,* the sum of rated values of all *personal* traumata.
5. *MTV—milieu traumatic value,* the sum of rated values of all *milieu* traumata.
6. *TLV—traumatic load value,* the sum of *PTV* plus *MTV.*

In addition to these six scores, three tentative global scores are suggested for the characterization of the whole content, as well as the personal and milieu sub-scores. These scores, which we may call the "global traumatic indices," are expressed in terms of *catastrophic units, CU.* The simple formula for the conversion of value scores into *CU*'s and its rationale are to be discussed later. However, in view of the tentative nature of the *CU* Index, the treatment of all data in the present article proceeds in terms of *raw scores* unless otherwise specified.

To test the proposed methodology, we have chosen 10 interviews[3] of Displaced Persons evenly subdivided in two distinct groups. We designate them accordingly as the Friendly Eastern Refugee group (*FER*) and the Concentration Camp group (*KZ*).

---

[3] Seven of the interviews dealt with in this study are published in Boder's book (8 & 9). These are Mr. Klüver (Kl—Julius Braun), Mrs. Prest (Pr—Paul), Mrs. Freich (Fr) Mrs. Kovitzka (Ko), Miss Lichtheim (Li), Mr. Gastfreund (Bs-Bassfreund), Mr. Matzner (Mt). The stories of Professor Odinoff (Od), Father Kharchenko (Kh), and Anna Braun (Br) are available on microfilm, and microcards (9). A limited quantity of mimeographed copies are also available, and may be obtained upon request from the author.

The Friendly Refugee group consists of five individuals of Christian faith, all escapees from Russia, or Russian-occupied territory. Two of them, Mrs. Prest (*Pr*) and Father Kharchenko (*Kh*), were citizens of the Baltic States which gained their independence from Russia after the first World War. During the German-Soviet pact the Baltic States were "reunited" with Russia, and Prest and Kharchenko saw in the subsequent German occupation of their country an act of liberation. Upon the defeat of the Germans on the Eastern front, Prest and Kharchenko were evacuated to Germany with the retreating German armies, and were either placed in refugee camps or billeted as workers with industries or farmers. They enjoyed reasonable freedom of movement and relatively tolerable conditions of existence. At no time was there any prohibition of the use of money or segregation of the sexes, nor were they restrained from mingling with the general German population. The other two persons in this Friendly Eastern Refugee group are Mr. Klüver (Julius Braun) and Mrs. Anna Braun. Both are Mennonites from the Ukraine, but not related to each other.[4] As Russian subjects, who for centuries continued to use the German language as language of worship and to a large extent as their language of communication, they were accepted by the Nazi government as Volksdeutsche, "Folk-Germans," i.e., as "Germans abroad," who in spite of their residence for more than a hundred years in Russia, had by Nazi definition never ceased to be Germans. Upon the occupation of the Ukraine by the German armies, the Mennonites were given permission and assistance to be evacuated to Germany. Like the Baltic refugees they enjoyed the privileges of "friendly" evacuees. Numerous Mennonites eagerly took advantage of this opportunity, first, as a means of escaping from under Soviet domination, and second, in the hope of joining eventually the Mennonite colonies on the American continent. Upon the defeat of Germany, they moved to the American Zone and refused "repatriation" to Russian-occupied territory.

A case by himself is presented by Professor Odinoff, the fifth in this group. He left Russia in 1919, retreating with the so-called "white (anti-Soviet) armies" and lived safely as an emigre in Paris until the German occupation in 1940. He then was interned as an "enemy alien," as were so many "white Russians," purely as a matter of military routine. However, apparently on account of his advanced age and severe chronic illness, he was finally released and lived safely in Paris up to the advent of liberation by the allies.

---

[4] At the time when the original translations were made, it appeared advisable to use "substitute names" for some of the interviewees. By some oversight, Klüver was given the name Braun, which now appears ill-advised since another extensive Mennonite story bears the real name of Mrs. Anna Braun.

In contrast to the group of Friendly Eastern Refugees stands the Concentration Camp (*KZ*) group, in this case represented by five Jews. The Nurenberg Laws, which amounted to a complete civic and legal disfranchisement of all "non-Aryans," made the persecution of the Jews mandatory. Cooperation of the civilian population in the implementation of the Nurenberg Laws was also considered mandatory. Four interviewees in this group, Miss Lichtheim (*Li*), Mrs. Kovitzka (*Ko*), Mr. Bassfreund (*Bs*), and Mr. Matzner (*Mt*), actually spent a number of years in concentration camps. The fifth, Mrs. Freich, like Professor Odinoff, was a refugee in Paris. She was arrested several times and sent to "distribution camps" in or near Paris, but was always set free on the strength of the fact that her husband was "voluntarily" working in a war plant, supplying wearing apparel to the German Army. However, two of her daughters were deported, and were not heard from afterwards, at least by the summer of 1946.

## C. THE TRAUMATIC INVENTORY AND THE ASSESSMENT OF TRAUMATIC FREQUENCIES

### 1. *Source and Method of Formulation of Themes*[5]

In the early stages of the project, attempts were made to adopt some existing designs used in the analyses of "personal documents." The "four wishes" of Thomas and Znaniecki (32) were definitely found wanting, which corroborates substantially some of the views expressed by Blumer (5) and Allport (1). Moreover, *The Polish Peasant* deals with slowly evolving social processes embracing an historical period protracted over decades, while the *DP* material deals with rapidly evolving catastrophical episodes.

More hopeful appeared Dollard's "criteria" (15). Hammond (17) and Uher (33), working under the direction of the present writer, zealously endeavored to use the Dollard "criteria" as a basis for content analysis. But the results appeared highly parsimonious. The reasons, besides some of those given by Allport (1), lie in the fact that a topical autobiography is not a life history. It lacks the protracted time factor during which personality development slowly evolves. Dollard himself had to admit the limitations of his method for topical biographies when he attempted to apply them to the Miller phantasies and poems (15, p. 263) used by Carl Young. However, the most important obstacle in using the "criteria" is the specific nature of the content of our interviews which Dollard could hardly foresee, nor intended to deal with. Dollard has definitely drawn his criteria as schemes devoted to the *acculturation process,* or what he called the *"group-plus-one"* hypothesis. In this connection Boder (8) writes:

---

[5] In addition to the references given in the text, it must be stated that in the design and more so in the appraisal of the methodology repeated cognizance was taken of the works of the following authors: Gordon Allport (1), Berelson (4), Blumer (5), Eron (15a, 15b, 15c), Fearing (16, 29), Lasswell, Leites *et al.* (18), Lazarsfeld (19), Merton, Lazarsfeld *et al.* (21), Parsons, Shils *et al.* (25), Ruesch and Bateson (27), and White (35a, 35b), as well as of the collective work usually designated as "The American Soldier" (31).

> Many of the recorded personal documents (discussed here) give precisely this picture in reverse—the . . . cutting down of a human being to fit into concentration and annihilation camps. To describe this process I had to make use of the antonym to the term *acculturation,* for what is dealt with here is an unprecedented and planned *deculturation* of personality on a mass scale.

The criteria for the description of a *deculturation* process overlap but *in part* with the acculturation criteria even if the latter were expressed in negative quantities, or antonym terminology. Like Thomas Znaniecki's *four wishes,* Dollard's criteria obviously apply to events extensively protracted in time. It is the rapid shock-like eruption of deculturating forces striking unexpectedly with unprecedented impact from and in a multiplicity of directions which furnish in our observation the main manifest attributes of catastrophes. It was therefore decided to derive the method of analysis from the historically and culturally almost unprecedented contents of the interviews themselves.

The reading of the material convinces one that to a great extent the interviews belong to the category of *special pleadings* (Allport, 1), although not exactly in the sense as specified by Allport. People "who have seen better times" and then find themselves as waifs and wards of charity; men and women who either have nowhere to go or don't want to go where in the eyes of the powers to be they supposedly belong (the *FER* who refused repatriation or the *KZ*'s expecting "disposition") feel often, strangely enough, a need to "apologize" for their precarious status, and do it in terms of traumata either personally experienced or eluded in flight. Traumatization is the common denominator of all the catastrophe interviews whether collected after a man-made holocaust or a nature-made Kansas City flood. Consequently, a Traumatic Inventory appeared a most appropriate tool for content analysis of such narratives.

The Traumatic Inventory in its present form consists of 46 themes (Berelson, 4; Child *et al.,* 13; Murray, 24) conceptualized by isolating discriminatory traumatic episodes as encountered by reading, or listening to more than 70 interviews. They cover only the plights of civilians (all interviewees were civilians) who underwent hardships either as a result of flight from discriminatory persecutions (the *FER* group) or were subjected to discriminatory persecution as a consequence of the Nurenberg Laws whether in letter or spirit (the *KZ* group). Sufferings from direct exigencies of military operations, such as bombings, rationing, housing shortage, etc., were not included in the themes, and were not coded in the analysis.

One must state here that the interviewees themselves seem to have taken the war as such in reasonable stride, but they considered themselves not "ordinary" war victims, such as would be the case of combat personnel or the general populations of invaded cities. The *DP*'s of 1946, sheltered in the *UNRRA* camps, christians and Jews alike, considered themselves victims of *discriminatory* persecution. They appraised their suffering not as exigencies of war in the conventional military sense, but as manifestations of organized hostility systematically directed against certain sectors of the civilian population, and based on political, ideological, economic, and religious idiosyncrasies.

Since the traumatic themes were to undergo a rating procedure by a group of judges, their wording had to be necessarily brief. In most cases, the themes present descriptions of generalized situations, rather than psychological or sociological abstractions (Appendix I). Their wording was aimed at arousing sufficiently differentiated states of empathic experiences as a precondition for effective rating. However, for the use in coding, the contents of the rated themes were extensively expanded, and in addition, a glossary was compiled to help the content analyst to find more readily the theme or themes which may cover a particular traumatic event.

### 2. *The Socio-Psychological Categories*

The classification of the traumata into the eight categories, such as socio-economic, cultural-affective, medical-organic, labor, bodily violence, hygiene, transport, and food (Tables 5, 6, 7), was undertaken as an afterthought, first simply as an additional help in locating a trauma in the Traumatic Inventory. The categories as they stand at present do not offer sharp delineations between the various psychosociological areas, but they reflect at least in most cases, the *predominant* nature of the trauma. Judging from the suggestive nature of the results, one could reasonably expect that a further refinement of the category classification, not only of the main themes, but of the subthemes within them as well, may reward itself in increased illuminating significance of the results.

### 3. *The Experiential Classification of Traumata*

In recent years psychology and psychiatry have come to use the dichotomous concepts of physical or somatic and psychic, or ideational without qualms or fears to be involved in the mind-body or any other monistic-dualistic controversies of yesterday. The traumatic themes once empirically formulated according to the actual contents of the interviews permit a post-factum classification into *predominantly* physical (including physiological, or somatic) and into *predominantly* ideological (psychological, cultural, attitudinal) traumata. Out of the 46 themes of the Traumatic Inventory, 29 are classified as physio-traumatic and 17 as ideo-traumatic. We also use the same terminology in characterizing a whole interview or a set of group data by designating them as physio-traumatic or ideo-traumatic depending upon the predominant experiential category in the contents. In case both categories are about evenly represented we designate the data or the whole interview as *equi-traumatic.* Following this procedure we may possibly encounter an interview in which the personal score may classify it as ideo-traumatic, the milieu score as physio-traumatic, and the traumatic load possibly as equi-traumatic.

## D. The Coding Procedure

The details of the coding procedures are given in Appendix II, and in part in more specific terms in the "interpretative expansions" of the Traumatic Inventory. We shall limit ourselves at this point only to the most important and general principles.

### 1. *The Problem of Multiple Assessments and Omissions*

The design of the project sets forth the task of ascertaining the fact of occurrence and frequency of actual incidents of specific traumata in an interview, without assessing repeated references to one and the same occurrence. The analysts were confronted with the problems of avoiding double or multiple assessments of the self-same events, as well as of mistaking a new specific event for one already reported. Such procedure requires from the coder an intelligent penetration into the continuity of the interview. The task is especially difficult in our type of semi-nondirective narratives, nearly always laden with affectivity, disruptions of chronological continuity by flash-backs, and often covering within an hour or two of verbal reporting, periods of five years or more.

### 2. *The Problem of Communication and Reliability*

Still more difficult was the task of precise formulation of the rules of coding. The analysis therefore was undertaken by the *case-conference method* (Allport, 1). First the investigator with the assistance of two, and at times of three graduate students made a content analysis of the 10 stories and reformulated and expanded the number of basic themes of the previously composed Traumatic Inventory. This was possible at the early stages of the investigation because the evaluation of the themes had not yet been undertaken. Subsequently a different group of three other graduate students were trained *by doing* in the analysis of the same stories. It was at that stage that the decision was taken to classify all traumata reported by the interviewees into the two basic groups: the Personal traumata (*PT*) and the Milieu traumata (*MT*), viz., 3P, 7M, etc. However, several traumata which in substance designate "vicarious" experiences have been given "personal" classifications. So a compelled or unavoidable attendance at an execution is covered by trauma No. 34 and is scored 34P, while the execution itself is covered by trauma No. 33 and is coded 33M.

The following procedure was then adopted. Each story was first coded separately by each of the three scorers. As soon as that was done it was taken up in conference, with or without the participation of the principal investigator. The wording of the rules was kept fluid, but no great changes were made in substance. The wordings of the rules were repeatedly re-edited in order to give them greater stability, and precision. The 10 stories were then re-coded for the third time. This time the first conference data of three stories were checked against a new set of scores obtained by the principal investigator. The remaining seven stories were re-coded first individually and then in conference by the principal investigator and one of the members of the previous coding panel. The detailed rules which were finally formulated are given in Appendix II. We present here a brief illustration of the coding procedure (codable items are given in italics):

> In January, 1945, the lager was abandoned. We were *transported* to Shimberg. These were all small villages, stations, that received us, until we came to the lager *Bergen-Belsen*. En route we had not dozens, but *hundreds of deaths*. People who were *old and sick and weak, and unable to stand the march, would fall on the road. So we marched on and on* . . . and then one night *we arrived in Bergen-Belsen* (Boder 8, p. 214).
>
> *Coding:* (1) "We were transported. . . ." Imposed removal from one place to another, covered by trauma No. 1 of the TI. *We*—a personal trauma. Therefore, the coding symbol is 1P. (2) *Shimberg* is an intermediate stop-over (not a camp, not a prison)—not coded. (3) The trip from Shimberg to Belsen is a continuation of "We were transported . . ."—not coded. (4) *Bergen-Belsen* is a concentration camp and the interviewee was part of the *collective* target—coding 8P. (5) "En route . . ." *hundreds of deaths*. This was a prolonged march on foot akin to a death-march. It presents a *different aspect*, although of the

**self-same transport, coding 41P. (6) Deaths of "other people" as a consequence of the march, 32M. (7) "People who were old and sick . . ." —lack of sympathy for the sick and no hospital provisions, 18M. (8) "So we marched on and on . . ." restatement of a self-same fact previously coded—no coding. (9) "And we arrived in Bergen-Belsen"—reference to self-same fact previously coded—no coding.**

The coding procedure adopted in this project is by no means simple. It requires diligent training and practice on the part of the coders. There are numerous factors that could account for lack of reliability from one coder to the other and in the same coder, when he repeats the procedure. The reason lies in the *microthemic* approach, in the method of "atomic" rather than "molecular" (Berelson, 4; Boder, 6). Speigelman, Terwilliger, and Fearing (29), contrary to Kaplan and Goldsen, have failed to attain higher reliability of the coding judges by means of more "detailed rules of classification," and even propose that the opposite hypothesis should be tested. However, in the light of our experience it appears that the possibly paradoxical results of Spiegelman *et al.* may lie not in the fact of expansion of rules as such, but in the fact of *insufficient* expansion. Our experience definitely indicates that *sufficiently broad* expansions of coding rules, *experimentally designed and tested,* may lead to substantial reduction if not to removal of ambiguities of "rules of classification."

The pooling of diagnostic efforts of several individuals in order to increase the validity or "psychological meaningfulness" of an investigation is not entirely new in the arsenal of psychological methodology. With the limitations in time and resources required for the training of consistent coders the case-conference method appeared the best means for the attainment of valid results (Allport, 1; Reckless and Selling, 26; Cartwright and French, 12; Murray, 24).

### E. EVALUATION

It was obvious that the reported traumata widely differ in their possible impact, and that the evaluation of a narrative in terms solely of simple trauma frequency is bound to give an incomplete and at times a misleading picture. The *TLF* ratings for the interviews of Miss Lichtheim (*Li*) and Mr. Matzner (*Mt*) discussed in the results confirm most emphatically this thesis. It was decided therefore to devise a method of evaluation based on the relative traumatic impact of each trauma as inferred from the empathic ratings of a number of judges. It is clear of course that the relative empathy "values" of most if not all traumata have strong socio-cultural as well as transitory pressure determinants. Geographical distance between the location of the raters and the place of occurrence of the events, as well as the intervening time between the events and the rating experiments have to be considered *a priori* of crucial significance, at least until specific evidence to the contrary is made available.

However, if a sufficiently valid method be devised to establish the empathy values for a certain group of traumata (or for that matter also "elators," "satisfiers") first separately within several admittedly divergent cultures then a step nearer will have been taken in establishing a scientific measure of cultural divergence or what we would call *cultural barriers.* The significance of such data for cross-cultural understanding is of course obvious (Gardner Murphy, 23), but a set of established measures valid for a specific culture does not become practically or theoretically invalid solely on account of its definable limitations. Limitations of a tool are not fallacies as long as the claims for it do not surpass the delineated possibilities.

### 1. *The Rating Procedure*

The rating of the traumata was performed by an adaptation of the Q-technique (Stephenson, 30; Mowrer, 22).

The 46 themes which form the Traumatic Inventory were printed on separate cards, and a sufficient number of sets were mimeographed for the use of the raters, all graduate or undergraduate students in classes of psychology. Each rater was seated at a table, where beforehand was placed a package of the 46 trauma-cards together with an array of eight large "place cards," each one bearing in large print a single number, in the following order: 1, 4, 7, 12, 12, 7, 4, 1. These numbers were to indicate the number of cards to be placed by the rater in each pile below. The raters were instructed to read their instruction sheet while the proctor read it aloud and to refer to it as often as necessary. With some incidental variations the instructions read as follows:

> This is an experiment in empathy—that means a study of peoples' ability to put themselves in the other person's place. You are to arrange the cards in eight piles according to the severity of the traumata *as you feel they would affect you.* The single card you believe to contain the most severe trauma of them all should be placed at the extreme right. Conversely, the single card you believe to contain the least severe trauma should be placed at the extreme left. The cards in the remaining six piles must contain traumata which in your judgment are definitely more severe than any one in all the piles to the left, and less severe than any one in all the piles to the right. The number of cards you are to place in each pile varies and is printed on each of the eight large cards before you. Do not hurry; there is no time limit. And

> **inspect your classifications several times until you feel that you have done your best.**

After the instructions were read, but before the signal to start was given, questions referring to the clarification of the procedure were invited and answered by the proctor. After the rating was completed by the whole group, the results of each rater were entered on special form sheets and evaluated from 1 to 8 according to its rated position from the least severe to the most severe trauma.

In analyzing the rating returns it became apparent that some raters (judging from the haphazard arrangement of the cards) did not render the same degree of coöperation as most of the group. It was decided therefore to set up *post factum* a qualifying standard for the judges, by correlating the value ratings of each judge with the ratings of each and every one of the group, and to accept only the ratings of those judges whose judgments correlate positively at the 1 per cent level of significance with at least 50 per cent of the judges. Out of a total of 65 judges, 44 individuals satisfied this criterion.

### 2. *Evaluation of the Traumata*

As a first step, each theme was assigned a *value,* based on the mean of its rated position obtained from all qualified judges. The values so computed represented in round numbers a range from 1.5 to 6.85. Such a range, of course, gave an extremely narrow spread of values. It was therefore decided to apply to the traumata more discriminative numerical indices without disturbing their relative value order obtained from the ratings.

The reconversion of the obtained values to such more discriminatory indices was based on the following considerations: It appears justifiable to assume that the whole gamut of hedonic experiences forms a continuum from "extreme" gratification to extreme distress. This would coincide with Wundt's (36) postulates on feeling qualities, with the hedonic zero point located in the middle (Figure 1).

We may then assume that the whole roster of hedonic experiences will be distributed according to a unimodal curve (normal or skewed) with the theoretical mode at the point of the zero-hedonic intensity. All the traumatic themes will then fall by definition to the right side of the theoretical mode,

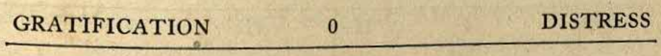


FIGURE 1
A SIMPLIFIED PRESENTATION OF WUNDT'S DIAGRAM OF FEELINGS

and if a normal distribution is postulated the grouping of the 46 themes in half-sigma areas should appear in approximate round numbers as given in Figure 2.

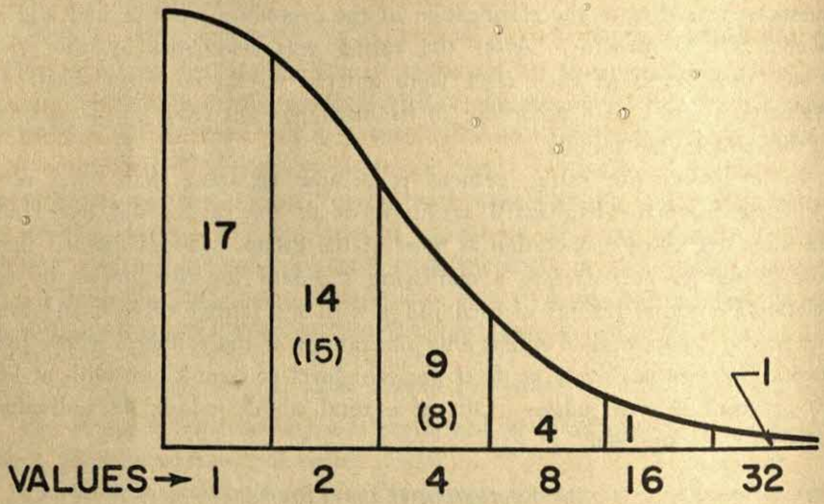


FIGURE 2
APPROXIMATE GROUPING OF THE 46 TRAUMATA INTO SIX HALF-SIGMA SECTORS OF THE AREA OF THE "HALF-BELL" CURVE: NUMBERS ALONG THE ABSCISSA DESIGNATE POSTULATED VALUES

The next step was to assign the values to the traumata (*T*'s) in each column. It was decided to use as value indices the numbers of a geometrical progression with the smallest integer above one as progression ratio. This gives a progression in terms of the powers of 2, i.e., 1, 2, 4, 8, 16, 32. In doing so we found, however, that the actual allocation of the previously rated traumata presented some difficulties with the columns at the second and third half-sigma. While eight cards definitely fell in the second column with the value of 2 and another eight cards in the third column with the value of 4, seven cards tied for either of the two positions. It was decided to place them all into the second column but with an intermediate value of 3. The final structure of the scale is given in Table 1. The values specifi-

TABLE 1
DISTRIBUTION OF TRAUMATA INTO HALF SIGMA AREAS (COLUMNS) OF THE RIGHT HALF OF THE NORMAL CURVE, WITH THEIR CORRESPONDING RATED VALUES

| Column, ½ sigma | I | II | III | IV | V | VI |
|---|---|---|---|---|---|---|
| No. of traumata | 17 | 15 | 8 | 4 | 1 | 1 |
| Traumatic values | 1 | 2 & 3 | 4 | 8 | 16 | 32 |

cally assigned to each trauma are given in the Traumatic Inventory (Appendix I), as well as in the Master Tables 2, 3, and 4. The number *one*, corresponding to the hypothetical evaluation of each of the 17 traumata in the first half-sigma sector has been designated as the *traumatic unit*, in terms of which the value scores must be read.

Once the traumata of an interview have been assessed according to the rules formulated for the content analysis, the evaluation of the traumatic content reduces itself to simple arithmetic. Each trauma is assigned its corresponding index value. The sum of these values will render the traumatic index value of all or any chosen group of traumata. As stated before we have then for each interview in addition to the three *frequency* scores: Personal trauma frequency (*PTF*), Milieu trauma frequency (*MTF*), traumatic load frequency (*TLF*), the three value scores: personal trauma value (*PTV*), milieu trauma value (*MTV*), and traumatic load value (*TLV*).

### 3. *The Traumatic Index in Terms of Catasprophe Units*

In order to reduce the numerical size of the value scores and possibly lend them a more intrinsic meaning we have proposed a method of conversion of value scores into catasprophe unit scores. As the unit we have chosen the number 8 which corresponds to the rated traumatic value of four themes in the traumatic index which appear most characteristic for man-made catastrophes. These themes are No. 16—perfidy and betrayal; No. 17—protracted roll calls lasting for hours often leading to physical exhaustion; No. 26—deterioration of or surgical interference with sex drive; and No. 31—"death marches." The traumatic index in terms of catastrophe units for the present is obtained by dividing the values by 8. A further refinement of this score may be possible by making it reflect the traumatic frequencies. For the time being we found the present frequency and value scores sufficient for the purposes of this investigation.

### 4. *Theme Coverage Score (Thcov.)*

The traumata may of course be used also for short-cut appraisals of catastrophic situations by ascertaining simply which of the traumata of the index occur at all in a single or in a set of reports. So it may be illuminating whether a certain biography contains or does not contain concentration camp situations, regardless of frequency of occurrence of such traumata. Such an assessment of theme coverage of a document may be illuminating in the understanding of an historical situation, but as we shall show in the section

"Results," would contribute but little to the appraisal of specific interviews and their intergroup or intragroup comparisons.

## F. RESULTS

The detailed results of assessment and evaluation of 10 interviews in terms of the Traumatic Inventory and the Evaluation Scale are presented in three Master Tables 2, 3, and 4 and in the scorogram (Figure 3). For the purpose of the discussion of the results in more general terms the data from the master tables have been condensed in Tables 6, 7, 8, and 9.

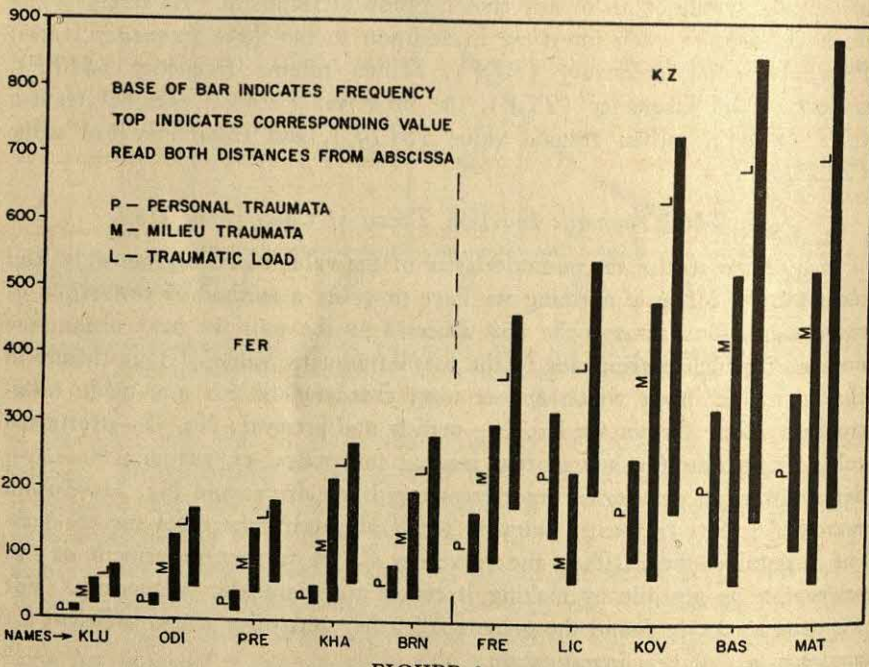


FIGURE 3
SCOROGRAM PRESENTING THE INDIVIDUAL TRAUMATIC INDICES FROM TABLES 2, 3, AND 4

Besides the abbreviations given above the master tables contain the following abbreviated headings:

*Inv. no.*—order number of the trauma in the Traumatic Inventory.

*Ps. Cl.*—classification of the trauma as *Ph,* predominantly physical, or *A,* predominantly attitudinal.

*Value*—rated evaluation of each trauma in traumatic units. On the bottom of each table we find:

*Themcov.*—theme coverage, the number of different themes utilized

**in coding of the interview regardless of its frequency. The value index of Themcov. is the sum of values of each theme composing the Themcov. count.**

### 1. *The Frequency Scores*

As stated before (p. 9), in outlining the procedure for the content analysis, the traumatic frequency gives the number of codable traumatic incidents reported in an interview. Duplicate references to the self-same events were not coded. In other words, if an interviewee reports three different incidents of deportation occurring at dfferent times, the deportation theme (T1 in the Traumatic Inventory) is assessed three times. On the other hand, if the narrator perseveres in his references to the same lost limb two or more times without adding to the incident new codable experiences, the lost limb trauma (T18) is coded once. Thus, the traumatic frequency is not a linguistic index of the number of references, as would be the case in a *DRQ* assessment (Dollard and Mowrer), but an index of reported occurrences.

Table 5 (Section 5a) gives the frequency scores, *personal, milieu* and *load* for each of the 10 subjects, taken from the Master Tables 2, 3, and 4. The means for the *KZ* group exceed the corresponding means of the *FER* group from two to four times. The differences are proportionally largest for the personal frequencies and least for the milieu frequencies. This reflects the fact to which we shall have repeatedly to refer that the *FER* records contain numerically less traumata in which the interviewees personally were the target than traumata inflicted upon "others." By fleeing from the enemy, the *FER*'s were spared the experiences many of which were part of the lot of the *KZ* group. Mainly, *none* of the *FER*'s were in concentration camps. The inspection of the data abundantly demonstrates the traumatic differences between the two groups. All three frequency scores of Mrs. Braun (Table 5a), who ranks *highest* among the *FER*'s in terms of the traumatic load, stand far below the corresponding scores of Mrs. Freich, who ranks *lowest* in the *KZ* group. Or, stated in quantitative terms, the *highest* individual *PTF, MTF,* and *TLF* scores of the *FER* group (Mrs. Braun) are 42, 38, 80 respectively, while the corresponding lowest scores for the *KZ* group (Mrs. Freich) are 82, 94, and 176.

A further inspection of Table 5a renders the following observations: In four out of five cases of the *FER* group, the personal frequency scores are *lower* than the corresponding milieu scores, while in four out of five cases of the *KZ* group the personal frequencies are *higher* than the milieu frequencies. One may infer that the *KZ* subjects have not only *more* to tell,

# TABLE 2

ASSESSMENT OF PERSONAL TRAUMATA: FREQUENCY (*PTF*) AND EVALUATION (*PTV*)

| | Inv. No. | Ps. Cl. | Val. | Friendly Eastern Refugee Group (FER) | | | | | | | | | | STot FER | | Concentration Camp Group (KZ) | | | | | | | | | | STot KZ | | Totals | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Mr. Kl. | | Md. Od. | | Mrs. Pr | | Fa.Kh. | | Mrs. Br. | | | | Mrs. Fr. | | Miss Li | | Mrs. Ko. | | Mr. Bs. | | Mr. Mt. | | | | | |
| | | | | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV |
| Socio-economic | 1 | P | 2 | 2 | 4 | 1 | 2 | — | — | 1 | 2 | 6 | 12 | 10 | 20 | 3 | 6 | 5 | 10 | 9 | 18 | 4 | 8 | 12 | 24 | 33 | 66 | 43 | 86 |
| | 2 | P | 1 | — | — | — | — | — | — | 2 | 2 | — | — | 2 | 2 | — | — | 1 | 1 | 1 | 1 | — | — | — | — | 2 | 2 | 4 | 4 |
| | 3 | P | 2 | — | — | 1 | 2 | 3 | 6 | — | — | 3 | 6 | 7 | 14 | 8 | 16 | 11 | 22 | 2 | 4 | 7 | 14 | 7 | 14 | 35 | 70 | 42 | 84 |
| | 4 | A | 1 | — | — | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 4 | 7 | 7 | 3 | 3 | 3 | 3 | 5 | 5 | 2 | 2 | 2 | 2 | 15 | 15 | 22 | 22 |
| | 5 | A | 1 | 1 | 1 | — | — | — | — | 4 | 4 | — | — | 5 | 5 | 3 | 3 | 2 | 2 | 5 | 5 | 4 | 4 | — | — | 14 | 14 | 19 | 19 |
| | 6 | P | 1 | 1 | 1 | 1 | 1 | — | — | 2 | 2 | 3 | 3 | 7 | 7 | — | — | 5 | 5 | 1 | 1 | 3 | 3 | 4 | 4 | 13 | 13 | 20 | 20 |
| Cultural-affective | 7 | A | 2 | 2 | 4 | — | — | 2 | 4 | 1 | 2 | 5 | 10 | 10 | 20 | 10 | 20 | 5 | 10 | 7 | 14 | 3 | 6 | 1 | 2 | 26 | 52 | 36 | 72 |
| | 8 | P | 1 | 1 | 1 | 6 | 6 | 1 | 1 | 3 | 3 | 2 | 2 | 13 | 13 | 17 | 17 | 14 | 14 | 11 | 11 | 7 | 7 | 19 | 19 | 68 | 68 | 81 | 81 |
| | 9 | A | 3 | 3 | 9 | 3 | 9 | 1 | 3 | 3 | 9 | 2 | 6 | 12 | 36 | 13 | 39 | 8 | 24 | 23 | 69 | 7 | 21 | 12 | 36 | 63 | 189 | 75 | 225 |
| | 10 | A | 4 | — | — | — | — | — | — | — | — | 1 | 4 | 1 | 4 | — | — | 1 | 4 | — | — | — | — | — | — | 1 | 4 | 2 | 8 |
| | 11 | A | 1 | — | — | 1 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 6 | 6 | 6 | 6 | 15 | 15 | 8 | 8 | 4 | 4 | 5 | 5 | 38 | 38 | 44 | 44 |
| | 12 | A | 1 | — | — | 1 | 1 | 3 | 3 | — | — | 2 | 2 | 6 | 6 | — | — | — | — | — | — | — | — | 1 | 1 | 1 | 1 | 7 | 7 |
| | 13 | A | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 4 | 8 | 4 | 8 | 4 | 8 |
| | 14 | A | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 1 | — | — | — | — | — | — | 1 | 1 | 1 | 1 |
| | 15 | A | 1 | — | — | 1 | 1 | — | — | — | — | — | — | 1 | 1 | 2 | 2 | 4 | 4 | 2 | 2 | — | — | 3 | 3 | 11 | 11 | 12 | 12 |
| | 16 | A | 8 | — | — | — | — | 3 | 24 | 1 | 8 | — | — | 4 | 32 | 2 | 16 | 1 | 8 | 2 | 16 | 2 | 16 | 1 | 8 | 8 | 64 | 12 | 96 |
| Medical-Organic | 17 | P | 8 | — | — | — | — | — | — | — | — | 1 | 8 | 1 | 8 | — | — | 4 | 32 | — | — | 1 | 8 | 4 | 32 | 9 | 72 | 10 | 80 |
| | 18 | P | 4 | — | — | 2 | 8 | — | — | — | — | — | — | 2 | 8 | 2 | 8 | 7 | 28 | 5 | 20 | 4 | 16 | 4 | 16 | 22 | 88 | 24 | 96 |
| | 19 | P | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 2 | 8 | — | — | 2 | 8 | — | — | 4 | 16 | 4 | 16 |
| | 20 | P | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 2 | — | — | 1 | 2 | 1 | 2 |
| | 21 | A | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 4 | — | — | 1 | 4 | 1 | 4 |
| | 22 | P | 16 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 23-27 | | | Not found in any of the present records | | | | | | | | | | | | | | | | | | | | | | | | | |
| Labor | 28 | P | 1 | 1 | 1 | — | — | — | — | 1 | 1 | 2 | 2 | 4 | 4 | — | — | 4 | 4 | 3 | 3 | 3 | 3 | 7 | 7 | 17 | 17 | 21 | 21 |
| | 29 | P | 4 | — | — | — | — | — | — | 1 | 4 | 1 | 4 | 2 | 8 | — | — | 2 | 8 | 1 | 4 | 1 | 4 | 2 | 8 | 6 | 24 | 8 | 32 |
| | 30 | | | Not found in any of the present records | | | | | | | | | | | | | | | | | | | | | | | | | |

TABLE 2 (*continued*)

| | Inv. | Ps. | | Friendly Eastern Refugee Group (FER) | | | | | | | | | | STot | | Concentration Camp Group (KZ) | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Mr. Kl. | | Mr. Od. | | Mrs. Pr. | | Fa.Kh. | | Mrs. Br. | | FER | | Mrs. Fr. | | Miss Li | | Mrs. Ko. | | Mr. Bs. | | Mr. Mt. | | STot KZ | | Totals | |
| | No. | Cl. | Val. | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV | PTF | PTV |
| Bodily Violence | 31 | P | 3 | — | — | — | — | — | — | — | — | 1 | 3 | 1 | 3 | 1 | 3 | 4 | 12 | — | — | 3 | 9 | 2 | 6 | 10 | 30 | 11 | 33 |
| | 32 | P | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 4 | 2 | 8 | 1 | 4 | 4 | 16 | 4 | 16 |
| | 33 | P | 32 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 32 | 1 | 32 | 3 | 96 | 3 | 96 | 8 | 256 | 8 | 256 |
| | 34 | A | 2 | — | — | 1 | 2 | — | — | 2 | 4 | — | — | 3 | 6 | — | — | 3 | 6 | 3 | 6 | 1 | 2 | 3 | 6 | 10 | 20 | 13 | 26 |
| | 35 | P | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 1 | 1 | 1 | 1 | 1 |
| | 36 | P | 2 | — | — | — | — | — | — | — | — | 2 | 4 | 2 | 4 | 1 | 2 | 4 | 8 | 1 | 2 | 1 | 2 | 3 | 6 | 10 | 20 | 12 | 24 |
| Hygiene | 37 | P | 1 | — | — | 1 | 1 | — | — | — | — | 2 | 2 | 3 | 3 | 3 | 3 | 7 | 7 | 1 | 1 | 6 | 6 | 3 | 3 | 20 | 20 | 23 | 23 |
| | 38 | P | 1 | — | — | — | — | — | — | — | — | — | — | — | — | 2 | 2 | 6 | 6 | — | — | 1 | 1 | — | — | 9 | 9 | 9 | 9 |
| | 39 | P | 3 | — | — | — | — | — | — | — | — | 1 | 3 | 1 | 3 | 3 | 9 | 3 | 9 | 1 | 3 | 3 | 9 | 3 | 9 | 13 | 39 | 14 | 42 |
| | 40 | P | 3 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 3 | 3 | 9 | — | — | 4 | 12 | 4 | 12 |
| Transport | 41 | P | 8 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 8 | 1 | 8 | 1 | 8 | 1 | 8 | 4 | 32 | 4 | 32 |
| | 42 | P | 3 | — | — | — | — | — | — | — | — | 1 | 3 | 1 | 3 | 1 | 3 | 1 | 3 | — | — | 1 | 3 | 2 | 6 | 5 | 15 | 6 | 18 |
| | 43 | A | 4 | — | — | — | — | — | — | — | — | 1 | 4 | 1 | 4 | 1 | 4 | — | — | — | — | 1 | 4 | 1 | 4 | 3 | 12 | 4 | 16 |
| | 44 | P | 3 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 3 | — | — | 1 | 3 | 1 | 3 |
| Food | 45 | P | 3 | — | — | 1 | 3 | — | — | 3 | 9 | 1 | 3 | 5 | 15 | 1 | 3 | 7 | 21 | 3 | 9 | 10 | 30 | 3 | 9 | 24 | 72 | 29 | 87 |
| | 46 | A | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 2 | — | — | 1 | 2 | 1 | 2 | 3 | 6 | 3 | 6 |
| Totals | | | | 11 | 21 | 21 | 38 | 16 | 44 | 27 | 53 | 42 | 86 | 117 | 242 | 82 | 165 | 133 | 317 | 98 | 249 | 94 | 322 | 115 | 349 | 522 | 1402 | 639 | 1644 |
| | | | | Kl | | Od | | Pr | | Kh | | Br | | STotFER | | Fr | | Li | | Ko | | Bs | | Mt | | STotKZ | | Totals | |
| Th.cov. | | | | 7 | 11 | 13 | 23 | 8 | 19 | 14 | 31 | 20 | 50 | 62 | 134 | 19 | 46 | 30 | 110 | 24 | 91 | 32 | 123 | 29 | 110 | 134 | 480 | 196 | 614 |

# TABLE 3

## ASSESSMENT OF MILIEU TRAUMATA: FREQUENCY (MTF) AND EVALUATION (MTV)

| | Inv. No. | Ps. Cl. | Val. | Friendly Eastern Refugee Group (FER) | | | | | | | | | | | | Concentration Camp Group (KZ) | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Mr. Kl. | | Mr. Od. | | Mrs. Pr. | | Fa.Kh. | | Mrs. Br. | | STot FER | | Mrs. Fr. | | Miss Li | | Mrs. Ko. | | Mr. Bs. | | Mr. Mt. | | STot KZ | | Totals | |
| | | | | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV |
| Socio-economic | 1 | P | 2 | 4 | 8 | 1 | 2 | 3 | 6 | 3 | 6 | 4 | 8 | 15 | 30 | 11 | 22 | 12 | 24 | 9 | 18 | 9 | 18 | 9 | 18 | 50 | 100 | 65 | 130 |
| | 2 | P | 1 | — | — | — | — | — | — | 1 | 1 | — | — | 1 | 1 | — | — | — | — | 1 | 1 | — | — | — | — | 1 | 1 | 2 | 2 |
| | 3 | P | 2 | — | — | 3 | 6 | 3 | 6 | — | — | — | — | 6 | 12 | 7 | 14 | 1 | 2 | 2 | 4 | 1 | 2 | 1 | 2 | 12 | 24 | 18 | 36 |
| | 4 | A | 1 | — | — | — | — | 1 | 1 | 1 | 1 | 1 | 1 | 3 | 3 | — | — | 4 | 4 | — | — | 1 | 1 | — | — | 5 | 5 | 8 | 8 |
| | 5 | A | 1 | 3 | 3 | 4 | 4 | 3 | 3 | 2 | 2 | — | — | 12 | 12 | 1 | 1 | — | — | 2 | 2 | 2 | 2 | — | — | 5 | 5 | 17 | 17 |
| | 6 | P | 1 | — | — | — | — | 1 | 1 | 3 | 3 | 3 | 3 | 7 | 7 | 5 | 5 | — | — | 1 | 1 | 1 | 1 | 1 | 1 | 8 | 8 | 15 | 15 |
| Cultural-affective | 7 | A | 2 | 1 | 2 | 1 | 2 | — | — | — | — | 1 | 2 | 3 | 6 | — | — | — | — | — | — | 1 | 2 | 3 | 6 | 4 | 8 | 7 | 14 |
| | 8 | P | 1 | 6 | 6 | 6 | 6 | 11 | 11 | 3 | 3 | 7 | 7 | 33 | 33 | 13 | 13 | 11 | 11 | 10 | 10 | 12 | 12 | 8 | 8 | 54 | 54 | 87 | 87 |
| | 9 | A | 3 | — | — | 2 | 6 | 4 | 12 | 4 | 12 | 5 | 15 | 15 | 45 | 14 | 42 | 6 | 18 | 11 | 33 | 4 | 12 | 9 | 27 | 44 | 132 | 59 | 177 |
| | 10 | A | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 11 | A | 1 | 1 | 1 | 1 | 1 | 3 | 3 | 3 | 3 | — | — | 8 | 8 | 6 | 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 10 | 10 | 18 | 18 |
| | 12 | A | 1 | 3 | 3 | — | — | — | — | 2 | 2 | 1 | 1 | 6 | 6 | — | — | — | — | — | — | — | — | 4 | 4 | 4 | 4 | 10 | 10 |
| | 13 | A | 2 | — | — | — | — | 1 | 2 | — | — | 1 | 2 | 2 | 4 | — | — | — | — | 2 | 4 | 1 | 2 | — | — | 3 | 6 | 5 | 10 |
| | 14 | A | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 15 | A | 1 | — | — | — | — | — | — | — | — | — | — | — | — | 3 | 3 | — | — | 3 | 3 | 1 | 1 | 2 | 2 | 9 | 9 | 9 | 9 |
| | 16 | A | 8 | — | — | 3 | 24 | 5 | 40 | 2 | 16 | 5 | 40 | 15 | 120 | 13 | 104 | 7 | 56 | 6 | 48 | 4 | 32 | 3 | 24 | 33 | 264 | 48 | 384 |
| Medical-Organic | 17 | P | 8 | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 8 | — | — | — | — | — | — | 2 | 16 | 3 | 24 | 3 | 24 |
| | 18 | P | 4 | 1 | 4 | 1 | 4 | 1 | 4 | — | — | 3 | 12 | 6 | 24 | 4 | 16 | 6 | 24 | 5 | 20 | 6 | 24 | 5 | 20 | 26 | 104 | 32 | 128 |
| | 19 | P | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 2 | 8 | — | — | — | — | 2 | 8 | 2 | 8 |
| | 20 | P | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 21 | A | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 22 | P | 16 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 16 | — | — | 1 | 16 | 1 | 16 |
| | 23-27 | | | Not found in any of the present records | | | | | | | | | | | | | | | | | | | | | | | | | |
| Labor | 28 | P | 1 | 2 | 2 | 1 | 1 | 1 | 1 | 3 | 3 | 2 | 2 | 9 | 9 | 3 | 3 | 9 | 9 | 7 | 7 | 2 | 2 | 1 | 1 | 22 | 22 | 31 | 31 |
| | 29 | P | 4 | — | — | 1 | 4 | 2 | 8 | — | — | — | — | 3 | 12 | — | — | — | — | 1 | 4 | — | — | 1 | 4 | 2 | 8 | 5 | 20 |
| | 30 | | | Not found in any of the present records | | | | | | | | | | | | | | | | | | | | | | | | | |

TABLE 3 (*continued*)

| | | | | Friendly Eastern Refugee Group (FER) | | | | | | | | | | STot | | Concentration Camp Group (KZ) | | | | | | | | | | STot KZ | | Totals | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Inv. No. | Ps. Cl. | Val. | Mr. Kl. | | Mr. Od. | | Mrs. Pr. | | Fa.Kh. | | Mrs. Br. | | FER | | Mrs. Fr. | | Miss Li | | Mrs. Ko. | | Mr. Bs. | | Mr. Mt. | | | | | |
| | | | | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV | MTF | MTV |
| Bodily Violence | 31 | P | 3 | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 3 | — | — | — | — | 1 | 3 | 1 | 3 | 3 | 9 | 3 | 9 |
| | 32 | P | 4 | — | — | — | — | — | — | 1 | 4 | 1 | 4 | 2 | 8 | — | — | 1 | 4 | — | — | 1 | 4 | 1 | 4 | 3 | 12 | 5 | 20 |
| | 33 | P | 32 | 1 | 32 | 2 | 64 | 1 | 32 | 5 | 160 | 3 | 96 | 12 | 384 | 1 | 32 | 2 | 64 | 10 | 320 | 12 | 384 | 12 | 384 | 37 | 1184 | 49 | 1568 |
| | 34 | A | 2 | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 2 | — | — | — | — | — | — | — | — | 1 | 2 | 1 | 2 |
| | 35 | P | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 36 | P | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| Hygiene | 37 | P | 1 | — | — | — | — | — | — | — | — | — | — | — | — | 3 | 3 | 2 | 2 | — | — | 1 | 1 | — | — | 6 | 6 | 6 | 6 |
| | 38 | P | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 39 | P | 3 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 2 | 6 | — | — | 1 | 3 | — | — | 3 | 9 | 3 | 9 |
| | 40 | P | 3 | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 3 | — | — | — | — | 1 | 3 | 1 | 3 | 3 | 9 | 3 | 9 |
| Transport | 41 | P | 8 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 42 | P | 3 | — | — | — | — | 1 | 3 | — | — | — | — | 1 | 3 | 1 | 3 | — | — | — | — | — | — | — | — | 1 | 3 | 2 | 6 |
| | 43 | A | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | 44 | P | 3 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| Food | 45 | P | 3 | — | — | 2 | 6 | 2 | 6 | — | — | 1 | 3 | 5 | 15 | 5 | 15 | 1 | 3 | — | — | — | — | 1 | 3 | 7 | 21 | 12 | 36 |
| | 46 | A | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| Totals | | | | 22 | 61 | 28 | 130 | 43 | 139 | 33 | 216 | 38 | 196 | 164 | 742 | 94 | 298 | 65 | 228 | 73 | 484 | 64 | 526 | 66 | 531 | 362 | 2067 | 525 | 2809 |
| | | | | Kl | | Od | | Pr | | Kh | | Br | | STotFER | | Fr | | Li | | Ko | | Bs | | Mt | | STotKZ | | Totals | |
| Th.cov. | | | | 9 | 45 | 13 | 64 | 16 | 69 | 13 | 57 | 14 | 65 | 65 | 300 | 19 | 80 | 14 | 66 | 16 | 68 | 21 | 92 | 19 | 84 | 89 | 390 | 154 | 690 |

# TABLE 4

## FREQUENCY (TLF) AND EVALUATION (TLV) OF TRAUMATIC LOAD

| | Inv. No. | Ps. Cl. | Val. | Friendly Eastern Refugee Group (FER) | | | | | | | | | | STot | | Concentration Camp Group (KZ) | | | | | | | | | | STot KZ | | Totals | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Mr. Kl. | | Mr. Od. | | Mrs. Pr. | | Fa.Kh. | | Mrs. Br. | | FER | | Mrs. Fr. | | Miss Li | | Mrs. Ko. | | Mr. Bs. | | Mr. Mt. | | | | | |
| | | | | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV |
| Socio-economic | 1 | P | 2 | 6 | 12 | 2 | 4 | 3 | 6 | 4 | 8 | 10 | 20 | 25 | 50 | 14 | 28 | 17 | 34 | 18 | 36 | 13 | 26 | 21 | 42 | 83 | 166 | 108 | 216 |
| | 2 | P | 1 | — | — | — | — | — | — | 3 | 3 | — | — | 3 | 3 | — | — | 1 | 1 | 2 | 2 | — | — | — | — | 3 | 3 | 6 | 6 |
| | 3 | P | 2 | — | — | 4 | 8 | 6 | 12 | — | — | 3 | 6 | 13 | 26 | 15 | 30 | 12 | 24 | 4 | 8 | 8 | 16 | 8 | 16 | 47 | 94 | 60 | 120 |
| | 4 | A | 1 | — | — | 1 | 1 | 2 | 2 | 2 | 2 | 5 | 5 | 10 | 10 | 3 | 3 | 7 | 7 | 5 | 5 | 3 | 3 | 2 | 2 | 20 | 20 | 30 | 30 |
| | 5 | A | 1 | 4 | 4 | 4 | 4 | 3 | 3 | 6 | 6 | — | — | 17 | 17 | 4 | 4 | 2 | 2 | 7 | 7 | 6 | 6 | — | — | 19 | 19 | 36 | 36 |
| | 6 | P | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 5 | 5 | 6 | 6 | 14 | 14 | 5 | 5 | 5 | 5 | 2 | 2 | 4 | 4 | 5 | 5 | 21 | 21 | 35 | 35 |
| Cultural-affective | 7 | A | 2 | 3 | 6 | 1 | 2 | 2 | 4 | 1 | 2 | 6 | 12 | 13 | 26 | 10 | 20 | 5 | 10 | 7 | 14 | 4 | 8 | 4 | 8 | 30 | 60 | 43 | 86 |
| | 8 | P | 1 | 7 | 7 | 12 | 12 | 12 | 12 | 6 | 6 | 9 | 9 | 46 | 46 | 30 | 30 | 25 | 25 | 21 | 21 | 19 | 19 | 27 | 27 | 122 | 122 | 168 | 168 |
| | 9 | A | 3 | 3 | 9 | 5 | 15 | 5 | 15 | 7 | 21 | 7 | 21 | 27 | 81 | 27 | 81 | 14 | 42 | 34 | 102 | 11 | 33 | 21 | 63 | 107 | 321 | 134 | 402 |
| | 10 | A | 4 | — | — | — | — | — | — | — | — | 1 | 4 | 1 | 4 | — | — | 1 | 4 | — | — | — | — | — | — | 1 | 4 | 2 | 8 |
| | 11 | A | 1 | 1 | 1 | 2 | 2 | 5 | 5 | 5 | 5 | 1 | 1 | 14 | 14 | 12 | 12 | 16 | 16 | 9 | 9 | 5 | 5 | 6 | 6 | 48 | 48 | 62 | 62 |
| | 12 | A | 1 | 3 | 3 | 1 | 1 | 3 | 3 | 2 | 2 | 3 | 3 | 12 | 12 | — | — | — | — | — | — | — | — | 5 | 5 | 5 | 5 | 17 | 17 |
| | 13 | A | 2 | — | — | — | — | 1 | 2 | — | — | 1 | 2 | 2 | 4 | — | — | — | — | 2 | 4 | 1 | 2 | 4 | 8 | 7 | 14 | 9 | 18 |
| | 14 | A | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 1 | — | — | — | — | — | — | 1 | 1 | 1 | 1 |
| | 15 | A | 1 | — | — | 1 | 1 | — | — | — | — | — | — | 1 | 1 | 5 | 5 | 4 | 4 | 5 | 5 | 1 | 1 | 5 | 5 | 20 | 20 | 21 | 21 |
| | 16 | A | 8 | — | — | 3 | 24 | 8 | 64 | 3 | 24 | 5 | 40 | 19 | 152 | 15 | 120 | 8 | 64 | 8 | 64 | 6 | 48 | 4 | 32 | 41 | 328 | 60 | 480 |
| Medical-Organic | 17 | P | 8 | — | — | — | — | — | — | — | — | 1 | 8 | 1 | 8 | 1 | 8 | 4 | 32 | — | — | 1 | 8 | 6 | 48 | 12 | 96 | 13 | 104 |
| | 18 | P | 4 | 1 | 4 | 3 | 12 | 1 | 4 | — | — | 3 | 12 | 8 | 32 | 6 | 24 | 13 | 52 | 10 | 40 | 10 | 40 | 9 | 36 | 48 | 192 | 56 | 224 |
| | 19 | P | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 2 | 8 | 2 | 8 | 2 | 8 | — | — | 6 | 24 | 6 | 24 |
| | 20 | P | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 2 | — | — | 1 | 2 | 1 | 2 |
| | 21 | A | 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 4 | — | — | 1 | 4 | 1 | 4 |
| | 22 | P | 16 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 16 | — | — | 1 | 16 | 1 | 15 |
| | 23-27 | | | Not found in any of the present records | | | | | | | | | | | | | | | | | | | | | | | | | |
| Labor | 28 | P | 1 | 3 | 3 | 1 | 1 | 1 | 1 | 4 | 4 | 4 | 4 | 13 | 13 | 3 | 3 | 13 | 13 | 10 | 10 | 5 | 5 | 8 | 8 | 39 | 39 | 52 | 52 |
| | 29 | P | 4 | — | — | 1 | 4 | 2 | 8 | 1 | 4 | 1 | 4 | 5 | 20 | — | — | 2 | 8 | 2 | 8 | 1 | 4 | 3 | 12 | 8 | 32 | 13 | 52 |
| | 30 | | | Not found in any of the present records | | | | | | | | | | | | | | | | | | | | | | | | | |

TABLE 4 (*continued*)

| | Inv. | Ps. | | Friendly Eastern Refugee Group (FER) | | | | | | | | | | STot | | Concentration Camp Group (KZ) | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Mr. Kl. | | Mr. Od. | | Mrs. Pr. | | Fa.Kh. | | Mrs. Br. | | FER | | Mrs. Fr. | | Miss Li | | Mrs. Ko. | | Mr. Bs. | | Mr. Mt. | | STot KZ | | Totals | |
| | No. | Cl. | Val. | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV | TLF | TLV |
| Bodily Violence | 31 | P | 3 | — | — | — | — | — | — | — | — | 1 | 3 | 1 | 3 | 2 | 6 | 4 | 12 | — | — | 4 | 12 | 3 | 9 | 13 | 39 | 14 | 42 |
| | 32 | P | 4 | — | — | — | — | — | — | 1 | 4 | 1 | 4 | 2 | 8 | — | — | 1 | 4 | 1 | 4 | 3 | 12 | 2 | 8 | 7 | 28 | 9 | 36 |
| | 33 | P | 32 | 1 | 32 | 2 | 64 | 1 | 32 | 5 | 160 | 3 | 96 | 12 | 384 | 1 | 32 | 3 | 96 | 11 | 352 | 15 | 480 | 15 | 480 | 45 | 1440 | 57 | 1824 |
| | 34 | A | 2 | — | — | 1 | 2 | — | — | 2 | 4 | — | — | 3 | 6 | 1 | 2 | 3 | 6 | 3 | 6 | 1 | 2 | 3 | 6 | 11 | 22 | 14 | 28 |
| | 35 | P | 1 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 1 | 1 | 1 | 1 | 1 |
| | 36 | P | 2 | — | — | — | — | — | — | — | — | 2 | 4 | 2 | 4 | 1 | 2 | 4 | 8 | 1 | 2 | 1 | 2 | 3 | 6 | 10 | 20 | 12 | 24 |
| Hygiene | 37 | P | 1 | — | — | 1 | 1 | — | — | — | — | 2 | 2 | 3 | 3 | 6 | 6 | 9 | 9 | 1 | 1 | 7 | 7 | 3 | 3 | 26 | 26 | 29 | 29 |
| | 38 | P | 1 | — | — | — | — | — | — | — | — | — | — | — | — | 2 | 2 | 6 | 6 | — | — | 1 | 1 | — | — | 9 | 9 | 9 | 9 |
| | 39 | P | 3 | — | — | — | — | — | — | — | — | 1 | 3 | 1 | 3 | 3 | 9 | 5 | 15 | 1 | 3 | 4 | 12 | 3 | 9 | 16 | 48 | 17 | 51 |
| | 40 | P | 3 | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 3 | — | — | 1 | 3 | 4 | 12 | 1 | 3 | 7 | 21 | 7 | 21 |
| Transport | 41 | P | 8 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 8 | 1 | 8 | 1 | 8 | 1 | 8 | 4 | 32 | 4 | 32 |
| | 42 | P | 3 | — | — | — | — | 1 | 3 | — | — | 1 | 3 | 2 | 6 | 2 | 6 | 1 | 3 | — | — | 1 | 3 | 2 | 6 | 6 | 18 | 8 | 24 |
| | 43 | A | 4 | — | — | — | — | — | — | — | — | 1 | 4 | 1 | 4 | 1 | 4 | — | — | — | — | 1 | 4 | 1 | 4 | 3 | 12 | 4 | 16 |
| | 44 | P | 3 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 3 | — | — | 1 | 3 | 1 | 3 |
| Food | 45 | P | 3 | — | — | 3 | 9 | 2 | 6 | 3 | 9 | 2 | 6 | 10 | 30 | 6 | 18 | 8 | 24 | 3 | 9 | 10 | 30 | 4 | 12 | 31 | 93 | 41 | 123 |
| | 46 | A | 2 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | 1 | 2 | — | — | 1 | 2 | 1 | 2 | 3 | 6 | 3 | 6 |
| Totals | | | | 33 | 82 | 49 | 168 | 59 | 183 | 60 | 269 | 80 | 282 | 281 | 984 | 176 | 463 | 198 | 545 | 171 | 733 | 158 | 848 | 181 | 880 | 884 | 3469 | 1165 | 4453 |
| | | | | Kl | | Od | | Pr | | Kh | | Br | | STotFER | | Fr | | Li | | Ko | | Bs | | Mt | | STotKZ | | Totals | |
| Th.cov. | | | | 11 | 49 | 19 | 71 | 18 | 72 | 17 | 108 | 25 | 100 | 90 | 400 | 25 | 93 | 31 | 114 | 26 | 129 | 35 | 142 | 30 | 113 | 137 | 591 | 227 | 991 |

### TABLE 5

CONDENSATION OF MASTER TABLES 2, 3, AND 4 IN TERMS OF SIX SCORES FOR EACH INTERVIEWEE

| | Friendly Eastern Refugee Group (*FER*) | | | | | | | Concentration Camp Group (*KZ*) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Kl | Od | Pr | Kh | Br | STotI | Mean | Fr | Li | Ko | Bs | Mt | STotII | Mean | *FER:KZ* | (Total) |
| 5a—Grouping of frequency data | | | | | | | | | | | | | | | | |
| PTF | 11 | 21 | 16 | 27 | 42 | 117 | 23.4 | 82 | 133 | 98 | 94 | 115 | 522 | 104.4 | 1:4.3 | ( 639) |
| MTF | 22 | 28 | 43 | 33 | 38 | 164 | 32.8 | 94 | 65 | 73 | 64 | 66 | 362 | 74.4 | 1:2.2 | ( 526) |
| TLF | 33 | 49 | 59 | 60 | 80 | 281 | 56.2 | 176 | 198 | 171 | 158 | 181 | 884 | 176.8 | 1:3.1 | (1165) |
| 5b—Grouping of evaluation data | | | | | | | | | | | | | | | | |
| PTV | 21 | 38 | 44 | 53 | 86 | 242 | 48.4 | 165 | 317 | 249 | 322 | 349 | 1402 | 280.4 | 1:5.8 | (1644) |
| MTV | 61 | 130 | 139 | 216 | 196 | 742 | 148.4 | 298 | 228 | 484 | 526 | 531 | 2067 | 413.4 | 1:3.8 | (2809) |
| TLV | 82 | 168 | 183 | 269 | 282 | 984 | 196.8 | 463 | 545 | 733 | 848 | 880 | 3469 | 693.8 | 1:3.4 | (4453) |
| 5c—Juxtaposition of *personal* frequency and value scores | | | | | | | | | | | | | | | | |
| PTF | 11 | 21 | 16 | 27 | 42 | 117 | 23.4 | 82 | 133 | 98 | 94 | 115 | 522 | 104.4 | 1:4.3 | ( 639) |
| PTV | 21 | 38 | 44 | 53 | 86 | 242 | 48.4 | 165 | 317 | 249 | 322 | 349 | 1402 | 280.4 | 1:5.8 | (1644) |
| 5d—Juxtaposition of *milieu* frequency and value scores | | | | | | | | | | | | | | | | |
| MTF | 22 | 28 | 43 | 33 | 38 | 164 | 32.8 | 94 | 65 | 73 | 64 | 66 | 362 | 72.4 | 1:2.2 | ( 526) |
| MTV | 61 | 130 | 139 | 216 | 196 | 742 | 148.4 | 298 | 228 | 484 | 526 | 531 | 2067 | 413.4 | 1:2.8 | (2809) |
| 5e—Juxtaposition of traumatic *load* frequency and value scores | | | | | | | | | | | | | | | | |
| TLF | 33 | 49 | 59 | 60 | 80 | 281 | 56.2 | 176 | 198 | 171 | 158 | 181 | 884 | 176.8 | 1:3.1 | (1165) |
| TLV | 82 | 168 | 183 | 269 | 282 | 984 | 196.8 | 463 | 545 | 733 | 848 | 880 | 3469 | 693.8 | 1:3.6 | (4453) |
| 5f—Scores (tentative) in terms of catastrophe units | | | | | | | | | | | | | | | | |
| PCU | 2.6 | 4.8 | 5.5 | 6.6 | 10.8 | 30.2 | 6.04 | 20.6 | 39.6 | 31.1 | 40.2 | 43.6 | 175.2 | 35.04 | 1:5.8 | ( 205.5) |
| MCU | 7.6 | 16.2 | 17.4 | 27.0 | 24.5 | 92.8 | 18.56 | 37.3 | 28.5 | 60.5 | 65.8 | 66.4 | 258.4 | 51.68 | 1:2.8 | ( 351.1) |
| LCU | 10.2 | 21.0 | 22.9 | 33.6 | 35.3 | 123.0 | 24.60 | 57.9 | 68.1 | 91.6 | 106.0 | 110.0 | 433.6 | 86.72 | 1:3.6 | ( 556.6) |
| Pages | 14 | 23 | 25 | 21 | 39 | 122 | | 40 | 40 | 32 | 40 | 36 | 188 | | | 310 |

but have more to tell about *themselves.* The two exceptions are represented by Mrs. Braun and Mrs. Freich, one in each group who form the "middle pair" of the tables and represent, in terms of Traumatic Values (see below) the *most* traumatized of the *FER* group and the *least* traumatized of the *KZ* group respectively. On the bases of these data it should be permissible to state that the interviews of the *FER*'s are predominantly *milieu-oriented,* while those of the *KZ* are predominantly *ego-oriented.* This *milieu-orientation* of the *FER* group vs. the *ego-orientation* of the *KZ* group becomes still more apparent from another set of data (Table 5). The total *MTF* for the two groups together (526) is by 113 traumata less than their total *PTF* (639). However, the *MTF* score for the Eastern Refugees alone shows an increase of 47 traumata (from 117 to 164), or 39 per cent, while the *KZ* group shows a frequency *decrease* of 160 traumata (from 522 to 362) amounting to a decrease of 30 per cent for the same two variables. The psychotherapeutic inferences about the difference in "object orientation" of the two groups should not be overlooked, but appears but of limited significance. The cause for the difference appears predominantly *situational;* as repeatedly stated, the *FER* tell about a milieu from which they have managed to escape, the *KZ* narrate about personal traumata experienced amidst a catastrophe.

The traumatic load proportions for the frequency data (*TLF*) of the two groups take an intermediate position between the *PTF* and *MTF* scores, which corroborates the general observation that pooled data are prone to mask to an appreciable extent the variabilities of the components (Boder, 7; Manzer, 20).

### 2. *The Value Scores*

Table 5b presents the value scores in terms of *personal, milieu* and *load* scores for each interview, again as taken from the master tables. Again the division between the two groups of *DP*'s is clearcut, with the means for the *KZ* group being from three to nearly six times as large as the corresponding means for the *FER* group.

The rank order of the *FER* cases in terms of *TLV*'s has not changed in comparison with their rank order in terms of *TLF*'s. However, the differences within the *TLF* scores as compared with the differences within the *TLV* scores are in several cases far from being proportional. Thus the difference in the *TLF*'s for Braun and Kharchenko amounts to 20 traumata in the direction of Braun (Table 5e), while the difference between their *TLV*'s amounts only to 13 value points (units) in the same direction. Again, the difference between the *TLF*'s for Prest and Kharchenko is only one single trauma in the direction of Kharchenko, while the difference between the *values* amounts to 86 value points in the same direction. It is obvious that the interview with Kharchenko contains several traumata of very high if not extremely high value ratings. By finding his *personal* and *milieu* scores, both in Tables 5c and 5d, we observe that while his *PTF* of 27 renders a value of 53, his milieu traumata with an *MTF* of only 33, corresponding to an increase of 22 per cent, give a traumatic value of 216 or an increase of 232 per cent.

Different insights are rendered by the evaluation data for the *KZ* group. Here their rank order according to *TLF*'s (Table 5a) is cardinally reshuffled by the value scores. Lichtheim, with the highest *TLF* is next to the lowest in her group in terms of *TLV*'s, while Bassfreund, with the lowest *TLF* in the group, gains the second from top rank in terms of *TLV*. The shift in rank for Mrs. Freich goes from the median position in her group in terms of *TLF* to the lowest intragroup position with a score difference of 82 value units below Lichtheim, her nearest group mate in rank. This fact appears to give ample confirmation to the discriminatory power of the value ratings.

In nine of the 10 cases the milieu value scores (*MTV*) are higher than the *PTV*'s. The exception is presented by the Lichtheim interview (Boder, 8; "Words hurt too much"). For this case one may state that any experienced reader of case histories or therapeutic protocols will not fail to discover the great preoccupation of this interviewee with her personal problems, and with her subjective evaluations and introspections. One cannot help observing in the narrative of this young woman an egocentricity bordering on narcissism. Objectively speaking, conditions of her concentration camp career, heart-rending as they are, were materially less grave in degree of gravity from those of her group mates. Although her final months of captivity were spent in the infamous extermination camp of Bergen-Belsen, the first years of her confinement were spent, according to her report, under more tolerable conditions of labor service in a textile plant with a group of but a few hundred young women apparently selected for their stamina, fitness for a clean appearance, and possibly intelligence, and education.

As we shall endeavor to show in the later discussion (p. 35), the proportions of the various categories of traumatic themes covered by an interviewee may serve as a lead for the diagnosis of the presence and intensity of deculturation symptoms in the personality.

The proportions of the totals for the two groups in terms of value scores have increased both for the *personal* and *milieu* values on the side of the *KZ* group. On the other hand, no appreciable difference is observable in that direction in the proportions of the load scores. This again confirms the observation about variabilities being suppressed by "lumping" together of data and could undoubtedly be remedied by a mathematical readjustment of the scale values without disturbing the rank classification of the traumata as obtained in the rating procedure. However, the factual analysis of the interviews in terms of *our* ethical and hedonic standards makes it apparent that the use of an evaluation scale in conjunction with the frequency scores gives the appraisal of the interviews a significantly greater psychological and sociological meaning than could the frequency scores alone.

### 3. *The Classification of Scores by Categories*

Tables 6, 7, and 8 give the scores of the 10 interviews in classification by socio-psychological categories. The model for these classifications is of course the Spranger typology or more directly, the Allport and Vernon test (2). As stated before, the classifications are not sharply drawn and were originally attempted as an aid for coding of the contents. However, the obtained data are suggestive, and appear useful in the differential appraisal of group data as well as of individual records. A glance inspection of the tables gives a picture of the two groups in terms of *absence* scores, designated by a dash (—). The proportion of absence scores all weigh in favor of the *FER* group, and are as follows: *PTF*'s—16:1, *MTF*'s—12:1, and *TLF*'s—8:0.

We shall discuss the category scores under five headings.

*a. The category classification of personal traumata.* Table 6a reveals that the two highest traumatic frequencies in terms of sub-totals for each group belong to the cultural-affective and socio-economic categories respectively. Individually, in nine out of 10 cases, the cultural-affective frequencies excel the corresponding socio-economic scores. In the tenth case (*Kh*) they are equal. Moreover, the smaller of the two top-frequency totals (soc.-econ.—112) differs from the next ranking score (hygiene—46) by 76 traumata, or 59 per cent.

However, we gain a distinctly different picture and obviously of greater psychological validity if we turn to Table 6b which gives the corresponding *traumatic values* of the frequency entries. The sub-totals for the cultural-affective category are still the highest for both groups, but their ratios to other categories show in several instances marked changes. Still greater is the divergence of the value scores from the frequency scores in terms of rank and intergroup proportions in the cases of other categories. Thus while the socio-economic category for the *FER* group still retains its second from

TABLE 6a

CONDENSED TABULATION OF PERSONAL TRAUMA-FREQUENCIES (*PTF*) BY MAIN CATEGORIES

| | Friendly Eastern Refugee Group (*FER*) | | | | | | Concentration Camp Group (*KZ*) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Category | Kl | Od | Pr | Kh | Br | STotI | Fr | Li | Ko | Bs | Mt | STotII | *FER*:*KZ* | Total |
| Soc.-Econ. | 4 | 4 | 4 | 10 | 16 | 38 | 17 | 27 | 23 | 20 | 25 | 112 | 1: 2.9 | ( 150) |
| Cult-Aff. | 6 | 12 | 12 | 10 | 13 | 53 | 50 | 49 | 53 | 23 | 46 | 221 | 1: 4.2 | ( 274) |
| Med-Org. | — | 2 | — | — | 1 | 3 | 2 | 13 | 5 | 9 | 8 | 37 | 1:12.3 | ( 40) |
| Labor | 1 | — | — | 2 | 3 | 6 | — | 6 | 4 | 4 | 9 | 23 | 1: 3.8 | ( 29) |
| Bod. Violence | — | 1 | — | 2 | 3 | 6 | 2 | 12 | 6 | 10 | 13 | 43 | 1: 7.2 | ( 49) |
| Hygiene | — | 1 | — | — | 3 | 4 | 8 | 16 | 3 | 13 | 6 | 46 | 1:11.5 | ( 50) |
| Transport | — | — | — | — | 2 | 2 | 2 | 2 | 1 | 4 | 4 | 13 | 1: 6.5 | ( 15) |
| Food | — | 1 | — | 3 | 1 | 5 | 1 | 8 | 3 | 11 | 4 | 27 | 1: 5.4 | ( 32) |
| Totals | 11 | 21 | 16 | 27 | 42 | 117 | 82 | 133 | 98 | 94 | 115 | 522 | 1: 4.5 | ( 639) |

TABLE 6b

CONDENSED TABULATION OF PERSONAL TRAUMA-VALUES (*PTV*) BY MAIN CATEGORIES

| Category | Kl | Od | Pr | Kh | Br | STotI | Fr | Li | Ko | Bs | Mt | STotII | *FER*:*KZ* | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Soc.-Econ. | 6 | 6 | 7 | 11 | 25 | 55 | 28 | 43 | 34 | 31 | 44 | 180 | 1: 3.3 | ( 235) |
| Cult-Aff. | 14 | 18 | 37 | 24 | 25 | 118 | 100 | 80 | 120 | 54 | 82 | 436 | 1: 3.7 | ( 554) |
| Med-Org. | — | 8 | — | — | 8 | 16 | 8 | 68 | 20 | 38 | 48 | 182 | 1:11.3 | ( 198) |
| Labor | 1 | — | — | 5 | 6 | 12 | — | 12 | 7 | 7 | 15 | 41 | 1: 3.4 | ( 53) |
| Bod. Violence | — | 2 | — | 4 | 7 | 13 | 5 | 58 | 44 | 117 | 119 | 343 | 1:26.2 | ( 356) |
| Hygiene | — | 1 | — | — | 5 | 6 | 14 | 22 | 7 | 25 | 12 | 80 | 1:13.3 | ( 86) |
| Transport | — | — | — | — | 7 | 7 | 7 | 11 | 8 | 18 | 18 | 62 | 1: 8.9 | ( 69) |
| Food | — | 3 | — | 9 | 3 | 15 | 3 | 23 | 9 | 32 | 11 | 78 | 1: 5.2 | ( 93) |
| Totals | 21 | 38 | 44 | 53 | 86 | 242 | 165 | 317 | 249 | 322 | 349 | 1402 | 1: 5.8 | (1644) |
| P.Tr.Inx. (CU) | 2.6 | 4.8 | 5.5 | 6.6 | 10.8 | 30.2 | 20.6 | 39.6 | 31.1 | 40.2 | 43.6 | 175.2 | 1: 5.8 | ( 205.5) |

### TABLE 7a
CONDENSED TABULATION OF MILIEU TRAUMA-FREQUENCIES (*MTF*) BY MAIN CATEGORIES

| | Friendly Eastern Refugee Group (*FER*) | | | | | | Concentration Camp Group (*KZ*) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Category | Kl | Od | Pr | Kh | Br | STotI | Fr | Li | Ko | Bs | Mt | STotII | *FER*:*KZ* | Total |
| Soc.-Econ. | 7 | 8 | 11 | 10 | 8 | 44 | 24 | 17 | 15 | 14 | 11 | 81 | 1: 1.8 | ( 125) |
| Cult-Aff. | 11 | 13 | 24 | 14 | 20 | 82 | 49 | 25 | 33 | 24 | 30 | 161 | 1: 2.0 | ( 243) |
| Med-Org. | 1 | 1 | 1 | — | 3 | 6 | 5 | 6 | 7 | 7 | 7 | 32 | 1: 5.3 | ( 38) |
| Labor | 2 | 2 | 3 | 3 | 2 | 12 | 3 | 9 | 8 | 2 | 2 | 24 | 1: 2.0 | ( 36) |
| Bod. Violence | 1 | 2 | 1 | 6 | 4 | 14 | 3 | 3 | 10 | 14 | 14 | 44 | 1: 3.1 | ( 58) |
| Hygiene | — | — | — | — | — | — | 4 | 4 | — | 3 | 1 | 12 | 0:12.0 | ( 12) |
| Transport | — | — | 1 | — | — | 1 | 1 | — | — | — | — | 1 | 1: 1.0 | ( 2) |
| Food | — | 2 | 2 | — | 1 | 5 | 5 | 1 | — | — | 1 | 7 | 1: 1.4 | ( 12) |
| Totals | 22 | 28 | 43 | 33 | 38 | 164 | 94 | 65 | 73 | 64 | 66 | 362 | 1: 2.2 | ( 526) |

### TABLE 7b
CONDENSED TABULATION OF MILIEU TRAUMA-VALUES (*MTV*) BY MAIN CATEGORIES

| Category | Kl | Od | Pr | Kh | Br | STotI | Fr | Li | Ko | Bs | Mt | STotII | *FER*:*KZ* | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Soc.-Econ. | 11 | 12 | 17 | 13 | 12 | 65 | 42 | 30 | 26 | 24 | 21 | 143 | 1: 2.2 | ( 208) |
| Cult-Aff. | 12 | 39 | 68 | 36 | 67 | 222 | 168 | 86 | 99 | 62 | 72 | 487 | 1: 2.2 | ( 709) |
| Med-Org. | 4 | 4 | 4 | — | 12 | 24 | 24 | 24 | 28 | 40 | 36 | 152 | 1: 6.3 | ( 176) |
| Labor | 2 | 5 | 9 | 3 | 2 | 21 | 3 | 9 | 11 | 2 | 5 | 30 | 1: 2.4 | ( 51) |
| Bod. Violence | 32 | 64 | 32 | 164 | 100 | 392 | 37 | 68 | 320 | 391 | 391 | 1207 | 1: 4.1 | (1599) |
| Hygiene | — | — | — | — | — | 0 | 6 | 8 | — | 7 | 3 | 24 | 0:24.0 | ( 24) |
| Transport | — | — | 3 | — | — | 3 | 3 | — | — | — | — | 3 | 1: 1.0 | ( 6) |
| Food | — | 6 | 6 | — | 3 | 15 | 15 | 3 | — | — | 3 | 21 | 1: 1.4 | ( 36) |
| Totals | 61 | 130 | 139 | 216 | 196 | 742 | 298 | 228 | 484 | 526 | 531 | 2067 | 1: 2.8 | (2809) |
| M.Tr.Inx. (CU) | 7.6 | 16.2 | 17.4 | 27.0 | 24.5 | 92.8 | 37.3 | 28.5 | 60.5 | 65.8 | 66.4 | 258.4 | 1: 2.8 | ( 351.1) |

TABLE 8a

CONDENSED TABULATION OF TRAUMATIC LOAD FREQUENCIES (*TLF*) BY MAIN CATEGORIES

| | Friendly Eastern Refugee Group (*FER*) | | | | | | Concentration Camp Group (*KZ*) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Category | Kl | Od | Pr | Kh | Br | STotI | Fr | Li | Ko | Bs | Mt | STotII | *FER:KZ* | Total |
| Soc.-Econ. | 11 | 12 | 15 | 20 | 24 | 82 | 41 | 44 | 38 | 34 | 36 | 193 | 1: 2.3 | ( 275) |
| Cult-Aff. | 17 | 25 | 36 | 24 | 34 | 136 | 99 | 74 | 86 | 47 | 76 | 382 | 1: 2.8 | ( 518) |
| Med-Org. | 1 | 3 | 1 | — | 3 | 8 | 7 | 19 | 12 | 16 | 15 | 69 | 1: 8.6 | ( 77) |
| Labor | 3 | 2 | 3 | 5 | 5 | 18 | 3 | 15 | 12 | 6 | 11 | 47 | 1: 2.6 | ( 65) |
| Bod. Violence | 1 | 3 | 1 | 8 | 7 | 20 | 5 | 15 | 16 | 24 | 27 | 87 | 1: 4.4 | ( 107) |
| Hygiene | — | 1 | — | — | 3 | 4 | 12 | 20 | 3 | 16 | 7 | 58 | 1:14.5 | ( 62) |
| Transport | — | — | 1 | — | 2 | 3 | 3 | 2 | 1 | 4 | 4 | 14 | 1: 4.7 | ( 17) |
| Food | — | 3 | 2 | 3 | 2 | 10 | 6 | 9 | 3 | 11 | 5 | 34 | 1: 3.4 | ( 44) |
| Totals | 33 | 49 | 59 | 60 | 80 | 281 | 176 | 198 | 171 | 158 | 181 | 884 | 1: 3.1 | (1165) |

TABLE 8b

CONDENSED TABULATION OF TRAUMATIC LOAD VALUES (*TLV*) FROM TABLE 4 BY MAIN CATEGORIES

(Raw Scores)

| Category | Kl | Od | Pr | Kh | Br | STotI | Fr | Li | Ko | Bs | Mt | STotII | *FER:KZ* | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Soc.-Econ. | 17 | 18 | 24 | 24 | 37 | 120 | 70 | 73 | 60 | 55 | 65 | 323 | 1: 2.7 | ( 443) |
| Cult-Aff. | 26 | 57 | 105 | 60 | 100 | 348 | 268 | 166 | 219 | 116 | 154 | 923 | 1: 2.6 | (1271) |
| Med-Org. | 4 | 12 | 4 | — | 12 | 32 | 32 | 92 | 48 | 78 | 84 | 334 | 1:10.0 | ( 366) |
| Labor | 3 | 5 | 9 | 8 | 8 | 33 | 3 | 21 | 18 | 9 | 20 | 71 | 1: 2.2 | ( 104) |
| Bod. Violence | 32 | 66 | 32 | 168 | 107 | 405 | 42 | 126 | 364 | 508 | 510 | 1550 | 1: 3.9 | (1955) |
| Hygiene | — | 1 | — | — | 5 | 6 | 20 | 30 | 7 | 32 | 15 | 104 | 1:17.3 | ( 110) |
| Transport | — | — | 3 | — | 7 | 10 | 10 | 11 | 8 | 18 | 18 | 65 | 1: 6.5 | ( 75) |
| Food | — | 9 | 6 | 9 | 6 | 30 | 18 | 26 | 9 | 32 | 14 | 99 | 1: 3.3 | ( 129) |
| Totals | 82 | 168 | 183 | 269 | 282 | 984 | 463 | 545 | 733 | 848 | 880 | 3469 | 1: 3.6 | (4453) |
| L.Tr.Inx. (CU) | 10.2 | 21.0 | 22.9 | 33.6 | 35.3 | 123.0 | 57.9 | 68.1 | 91.6 | 106.0 | 110.0 | 433.6 | 1: 3.6 | ( 556.6) |

the top position, as well as an analogous significant quantitative distance from the next ranking category, the socio-economic score for the *KZ* group yields its second from the top position to the category of physical violence, and ties for third place with the medical-organic category. In other words, 112 socio-economic traumata (Table 7a) yield a value score of 180 ***traumatic points*** (Table 7b) while 43 traumata of ***bodily violence*** give a value score of 343 traumatic points. If such findings appear psychologically and culturally meaningful, then the proposed argument for evaluation of traumata and to some extent the presented methods of evaluation have proved their rationale.

The psychological validity or meaningfulness of the evaluation difference in the procedure is best illustrated by a breakdown of the data for the category of physical violence as ***personal*** traumata into separate themes. We lift these data from Table 2 giving together with the trauma number, the frequencies and computed values.

TABLE 9

| DP | Tr. No. 31 | | Tr. No. 32 | | Tr. No. 33 | | Tr. No. 34 | | Tr. No. 35 | | Tr. No. 36 | | Totals | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| groups | F | V | F | V | F | V | F | V | F | V | F | V | F | V | CU |
| *FER*'s | 1 | 3 | — | — | — | — | 3 | 6 | — | — | 2 | 4 | 6 | 13 | 1.6 |
| *KZ* group | 10 | 30 | 4 | 16 | 8 | 256 | 10 | 20 | 1 | 1 | 10 | 20 | 43 | 353 | 42.9 |

As we may readily see from Table 9, the *FER* group have no entries for three out of the six traumata that form the category of physical violence. Among these absent themes is trauma No. 32 with a value rating of 4 traumatic units and trauma No. 33, the only item in the scale with a value rating of 32 traumatic units.

Other aspects of Table 6 will be discussed in conjunction with the other two category tables.

*b. The classification of milieu traumata.* As in the category table for *PTF*'s the cultural-affective and socio-economic categories have the two respective top ranking positions also in the milieu scores. As stated before (p. 24): although the total for the two groups is by 113 traumata less for the "milieu" than for the "personal" records, the sub-total for the milieu frequency for the *FER* has increased from 117 to 164 traumata. This increase of 47 traumata is predominantly accounted for by the absolute increase in the cultural-affective category amounting to 29 traumata. On the other hand, the same cultural-affective category shows for the milieu scores of the *KZ* group a decrease of 60 traumata. However, an analysis of the evaluation scores renders rather different insights (Table 7b). Thus while for the *FER* the increase of 55 per cent of the frequency in the cultural-affective category brings an *increase* of 88 per cent in the value for the same entry, the *KZ* records show that in spite of a decrease of 60 traumata in the same category, the value score renders an increase of that score by 12 per cent. Obviously the separate theme components of the *FER* and the *KZ* group records, although belonging to the same category, are neither identical in content nor in rated values. The Master Tables (3 and 4) confirm this observation. The cause for the discrepancy lies in the fact that together with a decrease in frequency in 8 of the 10 themes in the cultural-affective category of the *KZ* milieu record, the frequency of trauma No. 16 which is conceptualized as "perfidy and betrayal" and traumatization of children, with an evaluated rating of eight traumatic units, shows an increase to an extent capable of not only maintaining but increasing the total traumatic value index for the whole cultural-affective category of the same record.

*c. The category classifications in the record of traumatic load.* The record of traumatic loads (Table 8) classified by categories confirms again the observation that although pooling of data may give a useful set of insights, such pooling is bound to obliterate greatly the ideographic characteristics of the individual components. This holds both for qualitative as well as for statistical conceptualizations. Our traumatic load entries represent simply the arithmetical sum of the *personal* and *milieu* scores both for frequencies and values. In reading the table one must keep in mind that the *TL* table does not reveal whether the entries have occurred in one or in both of the specific classifications (Table 6 and Table 7). The *absence entries* (—) cover only such trauma positions which have absence scores both in the personal and milieu record. The highest two proportions in the "load" value table giving the value ratios of the *FER* data to *KZ* data are for the item Hygiene, including sanitation with a ratio 1:17, and deficiency in medical care, 1:10. The proportion for the item "bodily violence," which for the "personal" records rendered the ratios of 1:7 for the *frequencies* and 1:26 for the *values,* is reduced in the load record to 1:4.4 and 1:3.9 respectively. The reason lies in the fact that although the *personal* record of *FER*'s contains but few incidents of physical violence, and those with relatively reduced traumatic impact (Table 6), their milieu references contain in that category a substantial number of extreme traumata. We may therefore consider the "personal" record as the presentation of direct personal experiences; the milieu record as the presentation of vicarious experiences or sources of fear and anxiety, and the traumatic load as a presentation of a state of events composed of a mosaic of traumatic potentialities and forming a global situational description of the catastrophe.

*d. The category tables taken together.* The individual records as given in the Master Tables (Tables 2, 3, and 4) as well as in the category tables permit a number of exploratory procedures. One must remember that the categories as well as the Traumatic Inventory were not *a priori* structures. They are a result of content analysis of spontaneously presented narratives by a methodology designated by Bain as Logico-Systematic analysis (Blumer, 5) or what we would call, following Robert Penn Warren (35), *a method of experimental reading.* In reading and listening to the wire recorded interviews we have endeavored to formulate a sufficient number of themes and sub-themes to cover any of the specific traumata found in the interviews. The number and content of the themes were experimentally chosen and have all their counterparts in the actual content of one or more interviews. The fact that the Traumatic Inventory contains a roster consisting of 10 themes classifiable as cultural-affective, and only two themes covering starvation is not due of course to the fact that cultural-affective needs have priority over demands for food, but that in retrospect and perspective, the cultural-affective needs have a greater diversity of components, a deprivation of which is prone to bear a traumatizing effect. The displaced persons were liberated and saved from physical extermination but their "present" status (at the time of the interview) as homeless, countryless, and destitute people was directly connected with traumata which had befallen them from the beginning of the catastrophe. We deal here not only with selectivity of memory determined by perceptual intensity or need-motives, but with selective communication behavior manifesting itself in conscious or intuitive acts of choice of "what to tell first" or what to tell at all. A marked deviation of a record of an individual from a generally established pattern of distribution of traumata by categories may then be construed as a lead either towards specificity of a situation or towards specific personal traumatic affinities. We take an illustration again from the food category. In the *FER* load records the food category ranks fifth in severity, while in the same record of the *KZ* group it ranks seven. Did the *KZ*'s have better food than the *FER*'s? The contents of the interviews refute the justification of such an assumption. The reason for the divergence in rank order lies in the fact that the *FER*'s had still *less* complaints about the conditions of medical care, hygiene, and transportation. On the other hand, the highest individual *PTF* ratings in the food category belong to *Li.* and *Bs.* (Tables 6a, 6b). These were the two youngest people in the group and were constantly assigned to hard work. It stands to reason that for them the deprivation of food was of greater significance and is therefore more vivid in retrospect than for the others.

Let us take another low frequency item which may appear paradoxical. We refer to the category of transportation. This item includes the whole variety of transportation incidents from travel in overcrowded conveyances for the *FER*'s to death marches for the *KZ*'s. In Table 7a we find only 13 entries for the *KZ* group out of a total of 522 traumata for all categories. To be sure, the value index is rather high: 62 traumatic units for the 13 transportation traumata. But the frequency and its rank positions (the lowest rank among the eight) still appear rather low. A plausible explanation lies in the assumption that had these people been submitted to further transportation ordeals they most probably would not have been among the survivors. In this connection one cannot help visualizing that in the assessment and evaluation of a larger number of catastrophe records it could appear possible to attain a quantitative connotation for the limits of human endurance under specific catastrophic conditions.

*e. Intergroup and intragroup comparisons of the category classifications.* We shall combine on the one hand the socio-economic and cultural-affective categories into a socio-cultural category as dealing with the destruction of the cultural framework of existence and on the other hand the categories of medico-organic and physical violence which deal with the somatic aspects of life, into a "somatic" category. Following Table 6 we find that for the *FER* group the combined socio-cultural scores give a total *PTF* of 91 with a *PTV* of 173; the combined "somatic" scores render a *PTF* of 9 with a value (*PTV*) of 39. The *KZ* group attains a socio-cultural *PTF* score of 333 with a *PTV* of 616, and a somatic *PTF* score of 80 with a *PTV* of 525. We observe that while the proportions of the frequency ratings for the *socio-cultural* traumata of the two groups are practically equal to the proportions of their value scores (about 1:2), the proportions for the somatic categories are greatly divergent: not only do the personal somatic frequencies of the two groups render a ratio of 1:9, but the ratio of the values climbs to 1:18.

Another point: within the *FER* group the ratio of socio-cultural *PTF*'s to the somatic *PTF*'s is 10:1; for the *KZ* group this ratio is 4:1. How could the difference in the latter two proportions be interpreted? Would that mean that the *KZ* interviews (or by inference the persons themselves) were less "culture-directed" than the *FER*'s? The absolute scores of course refute that. Nevertheless, the fact that the proportional distribution of the communication content itself shows this difference between the two sets of interviews is prone to suggest some psycho-sociological insights. It is not remote to hypothesize that the prolonged manipulation of human beings combined with constant threats to their lives in the form of violence and disease, coupled with a rather hopeless outlook for the future at the time of the interview may well have depressed their sensitivity toward cultural needs and deprivations. An interviewee is normally under a pressure of choice between events that crowd his memory and this choice wherever possible goes in the direction of the more dramatic, or higher priority of need, as evaluated by the narrator. The arsenal of recollections of the *KZ*'s were obviously more abundant in personal somatic experiences than the stock pile of memories of the *FER*'s.

It is however not excluded that a low proportion of culturo-economic items in an interview against a high proportion of somatic contents could serve as a lead to the discovery of the operation of *deculturation* factors. The term "deculturation" must be understood as referring to two different concepts: (*a*) the deculturated or deculturating environment (the verbs are used both in their transitive and passive connotations) and (*b*) deculturated personality. A deculturated environment such as a concentration camp, slums, lock-ups of police stations, bombed-out cities or any makeshift installation in substitution of standard conditions and attributes of existence is bound to evoke manifestations of subcultural behavior in its victims. On the other hand deculturation of personality manifests itself not in the physical submission but in the intellectual and affective acceptance of the materially and ethically deculturated mode of existence. The problems of diagnosis of deculturation and the technique of *reculturation* of personality stand far beyond their purely theoretical significance in a world which counts about twelve million displaced persons.

By its exterior manifestations and immediate effect catastrophe is predominantly an acute deculturating phenomenon, and its impact could possibly be measured by the depression of the level of subsequent needs and demands of its victims. However, the variation in the ratio of various psychological categories within an interview may be also influenced by factors emanating from the nature of a particular type of communication process (spoken interviews, letters, etc.). Like the *Language of Politics* (Lasswell, Leites, *et al.*, 18) the language of catastrophe has its specific characteristics. These will be discussed in the second article of this series dealing with the linguistics of catastrophe.

### 4. *The Experiential Classification of the Interviews*

The 46 themes of the Traumatic Inventory fall into two experiential groups: 29 themes rate as physio-traumatic and 17 as ideo-traumatic. The subclassification of the traumata into the two experiential categories gives for each group as well as for each individual interviewee six pairs of scores or 60 individual pairs for the 10 interviews. To illustrate: Mrs. Prest has in the *PTF*'s four physio-traumata and 12 ideo-traumata. This diagnoses her *PTF* score as (predominantly) ideo-traumatic. The corresponding *PTV*'s are 7 and 37 respectively, which puts her *PTV* also in the ideo-traumatic category. Making use of a statistical shortcut by which any two scores with a difference less than 10 per cent of their sum are considered equi-traumatic we find that the group scores of the *FER*'s are equi-traumatic for *PTF*'s and *PTV*'s and physiotraumatic (frequency and value) both for the group scores of the milieu and traumatic load. The *KZ*'s are physio-traumatic in all six group scores. The relative lesser impact of physical traumatization for the *FER* group who were not confined to concentration camps is again reflected in the emerging validity of the distinct diagnostic differences between the *FER* and *KZ* groups, which in turn appears to corroborate the validity of the presented methods.

However, we may go beyond differentiations between groups on the basis of the described score classifications, and attempt to discover in the discrepancy between the tendency of the score of a particular individual and the tendency of the collective score for his group, the operation of specific factors either of situational or personality nature. We shall limit ourselves to a few examples.

For the *FER* group which for both *personal* scores has been characterized as equitramautic, we return to Mrs. Prest, a school teacher, with an ideo-traumatic designation both for *PTF* and *PTV*. These diagnostic designations correspond to the facts of her story. She was personally comparatively less disturbed in the material arrangements of her life and profession than her group mates, but she was harassed with the requirements of abolition of teaching of religion, and "Sovietization" of textbooks. From the facts one may conclude that the predominance of ideotraumata in her record are predominantly due to situational and not personal factors.

We turn now to the *KZ* group. In the "personal" classification of this group we have two deviants from the physiotraumatic diagnosis for the group—Mrs. Freich and Mrs. Kovitzka. In the *PT* category, Mrs. Freich is equitraumatic as to frequency and ideotraumatic as to values. Again the "middle" rank position of Mrs. Freich emerges to the foreground. She lived in German-occupied Paris, not in Germany proper. She was taken twice to distribution camps presumably for deportation, but freed both times within less than 15 days. She was not separated from her family, but she tragically lost two daughters, and was harassed physically and ideologically by the impact of the Nurenberg laws. In her narrative she most often accompanies the report on physical traumata with cultural appraisals, with pleas for understanding and justice. The reasons for the deviation of her diagnostic designations from that of her group seem both *situational* in the form of a comparatively lesser traumatic environment than was the case for her group mates, and *personal* in terms of presence and preservation of sensitivity to the cultural aspects of events. The latter observation is corroborated by the fact that she is the only one among the 10 interviewees who has a deviate score also in the milieu category. There her *value* score (*MTV*) has an ideotraumatic qualification, although the *MTF* is, like for the rest of her group mates, physiotraumatic.

The other deviate case, that of Mrs. Kovitzka, in the *KZ* group stands by itself. In the personal classification she has an *ideotraumatic* designation for the frequencies (*PTF*) and an *equitraumatic* designation for the value scores (*PTV*). The traumatic load of her story ranks her as third from the top among the 10 subjects, and falls in the middle position among her *KZ* group mates. Like Mrs. Freich, she is introspective, and accompanies the reports of her personal sufferings with a great deal of cultural and ethical evaluations. But in addition she still appears to continue under the pressure of guilt feelings derived from having given up her baby, and in a manner that eventually led to the death of the child, and from having been the "cause" of brutal treatment of her husband by the Gestapo when he attempted to save her from deportation. Here the ideotraumatic and equitraumatic aspects of the two scores, *PTF* and *PTV* respectively, seem to be rooted predominantly in personality factors.

By way of perspectives, it will not be remote to suggest that in dealing with larger quantities of such records, as with other clinical records and case studies, the described procedure may lead to a development of a device akin to a physio-ideo quotient which could yield quantitative indices of the physiotraumatic or ideo-traumatic qualities of contents of case interviews, or for that matter, any specimen of communication for which the content analysis as to traumata (or elators) may appear illuminating.

### 5. *The Theme Coverage and Absence Scores*

A glance at any of the three Master Tables 2, 3, and 4 permits to observe practically by inspection the difference in density of traumatic entries between the portions covering the *FER*'s and the *KZ* group. The line *Thcov.* (theme coverage) on the bottom of each table gives the respective totals of such entries regardless of frequency for the Personal, Milieu, and Load classifications as 62, 65, and 90. The same group scores for the *KZ* are 134, 89, and 147. The absence scores for each group may be obtained by subtracting the *Thcov.* score from 230, the total number of possible entries for each group. We observe that the group differentiations between the *FER*'s and *KZ*'s are still preserved in the occurrence or trauma coverage scores.

The respective evaluation indices for the *Thcov.* scores, again in terms of *PTV, MTV,* and *TLV* are for the *FER* group 134, 300, and 400 traumatic units, while for the *KZ* group they come to 480, 390, and 591. We observe then that for group comparisons or qualitative appraisal of *situations,* the "single occurrence" or theme coverage scores and their corresponding value scores may render illuminating results. But such also may possibly be the case by ascertaining the number of individual themes regardless of frequency evoked in the coding of the records of each group. We find that for the *FER* group 29 out of 46 different themes of the Inventory were used, with a traumatic value of 105, while for the *KZ* group 40 themes were utilized with a traumatic value of 155.

Useful as these indices may appear for the short-cut appraisal of catastrophic situations they are by far insufficient for the assessment of individual interviews and the discovery of valid content differences between them. A scrutiny of the individual *Thcov.* scores with their values in all three tables shows that in assessing the fact of traumatic occurrences only, the rank positions of the individual interviewees become unduly reshuffled both in terms of "occurrence" and value and the interpersonal differences become obscured and frequently reversed.

### 6. *The Use of the Traumatic Indices in Case Presentations*

The traumatic frequencies in conjunction with the value scores, whether as given in the master tables or in the category tables, may of course be used for the purpose of reconstruction and conceptual presentation of individual interviews. For that purpose the themes covered by entries may be verbalized and their impact described in terms of frequencies and values. The category tables may appear more illuminating than the master tables in view of their conceptualized structure. But the master tables present the results with higher ideographic fidelity. Wanting to some extent are numerical averages for trauma frequencies and values which could be used as reference quantities for the diagnostic appraisal of individual interview scores.

We refer to "norms" obtained from a statistically reliable universe. However, the data for the presented 10 interviews could well serve as a *pro tem* yardstick at least in a "clinical" approach to the appraisal of the material and the people who presented it.

### 7. *The Length of Interviews in Relation to the Traumatic Indices*

The sole factor in determining the length of an interview (Table 5, last line) was the willingness of the interviewee "to talk." The total length of the five *FER* interviews (122 pages) is significantly less than the number of pages covering the *KZ* records (188 pages). However, the size of the numerical score in terms of frequencies and values cannot be considered a direct function of the length of the interview, since these data show substantial differences in intergroup as well as intragroup comparisons of interviews of nearly identical length.

### 8. *The Catastrophe Units (CU)*

The conversion of value scores into *CU* scores for all three categories, personal, milieu, and load, have been discussed above and the scores are presented in Table 5f. Their purpose so far is to translate the value scores into smaller numbers without altering their intergroup ratios.

## G. CONCLUSIONS

1. Ten topical autobiographies of displaced persons verbatim wire-recorded in camps and shelterhouses of Western Europe in the summer of 1946 were used to design a methodology for the assessment and evaluation of traumatic experiences caused by man-made catastrophe. The interviewees belong to two distinct groups: five are classified as "friendly eastern refugees" (*FER*), the other five are designated as the "concentration camp group" (*KZ*).

2. A method of content analysis has been designed to deal with the structure of catastrophe interviews in terms of (*a*) frequency indices consisting of the sum of reported traumatic incidents, and (*b*) evaluation indices designed to characterize the traumatic impact of these incidents.

3. The content analysis was performed with the use of a Traumatic Inventory consisting of 46 traumatic themes conceptualized on the bases of more than 70 original recordings of *DP* interviews (Appendix I).

4. The evaluation scale was constructed on the basis of a rating procedure applied to the themes of the Traumatic Inventory using the principles of Stephenson's Q-technique. Subsequently the ranked scores of the trau-

mata were arranged in six groups approximately proportional in size to the half sigma areas of the "half bell" distribution curve. To each group was assigned a hypothetical numerical index value progressively increasing (in terms of the powers of 2) with the distance of the corresponding half-sigma sector from the mode, which was rated as the theoretical "hedonic zero." The so devised traumatic evaluations of the items range from 1 to 32. The "value" score of an interview is then computed by adding the index values of the traumata assessed in the content analysis.

5. Taking cognizance of the targets of the reported traumata, the content analysis permits differentiation between three sub-classifications: (*a*) Personal traumata (*PT*), those which affected the interviewee as a direct target; (*b*) Milieu traumata (*MT*), such as affecting "other persons," considered by the interviewee as group mates; and (*c*) Traumatic load scores (*TL*) which represent the sum of the *PT*'s and *MT*'s and cover the total traumatic content of the interview. In terms of frequencies the content analysis yields the *PTF, MTF,* and *TLF.*

6. The computation of values for the three frequency scores render three value indices: (*a*) Personal Traumatic value (*PTV*); (*b*) Milieu Traumatic value (*MTV*); and (*c*) Traumatic Load value (*TLV*).

7. The frequency and value indices have shown sufficient sensitivity to permit intergroup as well as intragroup comparisons.

8. The breakup of the traumatic themes into eight psycho-sociological categories such as socio-economic, medical, physical violence, food, etc., permits a profile presentation of the scores for individuals as well as for groups and illuminates more specifically the various psycho-sociological aspects of particular catastrophic situations, as well as the degree of emphasis laid upon them by the various interviewees.

9. The measurements confirm qualitatively and quantitatively the differences in traumatization undergone by the two distinct samples of displaced persons.

10. The numerical prevalence of milieu traumata (*MTF*) in the *FER* group over their *PTF* score and the reversed picture for the *KZ*'s who show a higher *PTF* than *MTF* rating permits to qualify the group scores of *FER*'s as milieu-oriented, while the corresponding *KZ* scores appear ego-oriented.

11. The totals for the *milieu value* scores are higher than the respective totals for personal value score for both groups. This does not invalidate the conclusions presented in the preceding paragraph as to the milieu-orientation of the *FER*'s and ego-orientation of the *KZ*'s, but rather indicates that both the *FER*'s as well as the *KZ*'s, witnessed in their environment incidents of

such traumatic impact that had they happened to them personally they most probably would not have lived to be interviewed.

12. The comparison of the scores grouped by psycho-sociological categories indicates that the proportions of frequency and value scores for the two groups greatly vary from category to category and deviate from the proportions of the totals to various degrees. So while the totals for the "personal" scores of the *FER* and *KZ* groups show a ratio of 1:4.5 for the traumatic frequencies, and of 1:5.8 for the traumatic values, the separate proportions for physical violence are 1:7.2 and for hygiene 1:11.5 in terms of *frequencies*, and 1:26.2 and 1:13.3 respectively in terms of *values*.

13. The value scores not only supplement the frequency scores but definitely sharpen and often correct the appraisals obtained by the frequency scores alone. The proportions for physical violence referred to in the previous paragraph are an illustration in point. As a whole, the value ratings appear to present the data in a more meaningful and more valid perspective than were we to appraise an interview from the frequency scores alone.

14. The "Theme coverage" score (*Th.cov.*) equivalent to a count of the number of *different* main themes used in coding of an interview regardless of the frequency of occurrence of the diverse incidents fitting into the theme, may give some idea of the "global" situation dealt with in an interview in a kind of all-or-none fashion. Moreover, the value index of the "Theme coverage" score may enhance the general appraisal of the reported situation. However, an attempt to use the *Th.cov.* score for intergroup or intragroup comparisons is prone to produce intensely misleading results.

15. The subclassification of the themes forming the Traumatic Inventory into *physical* (including somatic) and *attitudinal* (ideological) traumata reveals that the group records of the *FER*'s for personal traumatization both in terms of frequency (*PTF*) and value (*PTV*) fall in the equitraumatic category, while their respective milieu and load group records qualify as physiotraumatic. For the *KZ* group all three sets of group records fall in the physiotraumatic classification. This partial difference in the experiential classification between the two groups again corroborates the observation that the *FER*'s report their personal experiences more often in terms of fears of what *could have happened* to them (ideational experiences) while the *KZ* group reports the traumatic events in terms of occurrences that have actually struck them in person (physical, somatic events). Both groups of interviewees contain cases of deviations of individual *experiential tendencies* from the corresponding tendencies demonstrated by the group "as a whole." These ideographic discrepancies seem to supply promising leads for the discovery of

specific situational or personal factors operative in the reported sequence of events.

16. The subclassification of the themes which permits to diagnose the content trends of the interviews as ego-oriented vs. milieu-oriented, as physio-traumatic (somo-traumatic) vs. ideo-traumatic, both in terms of frequency and value indices, provides leads toward the discovery and appraisal of deculturation factors operating in catastrophic situations, and the assessment of interindividual differences in terms of susceptibility to the deculturating impact of catastrophic events.

17. The method of appraisal of traumatic experiences by means of the "half-bell" distribution, and their evaluation in terms of numerical values geometrically increasing with their distance from the theoretical mode gives impetus to the proposal of a psycho-sociological discipline of scientific *hedonology* composed of two fields: *traumatology* and its counterpart *elatology*. This should lead to the design of a methodology for quantitative evaluation of the various facets of social and personal experiences and to the possible construction of a lexicon akin to the *Occupational Dictionary* (34) which however *in addition to a listing and classification of* traumata would accompany the individual items with an evaluation index experimentally obtained within our culture. A cross-cultural evaluation subsequently undertaken may lead to a numerical presentation of cultural similarities, cultural barriers, and cultural distances, between people and countries of different cultural heritage.

## H. APPENDIX I: THE TRAUMATIC INVENTORY

(The themes of the Traumatic Inventory are given here in the form as they were originally printed on the cards for the rating judges. The *interpretative expansions,* which in the laboratory copy accompany most of the themes, are not given here in complete form for reasons of economy of space; they are left in part with some items for purposes of illustration. However, the number of expansions (*Exp*) available for each theme is given in parenthesis, preceded by the scale value of the theme and its experiential classification (*Ph*—physical; *A*—attitudinal). Complete mimeographed copies of the Traumatic Inventory may be obtained by writing to the author.)

### I. SOCIO-ECONOMIC AND GEOGRAPHICAL

1. Brutal and abrupt removal of a person from most environmental stimuli which have formed the conditioning framework of his everyday life. (2Ph;2Exp)

*Interpretative Expansion—Relocation.* (*a*) Flight in the path of war. (*b*) Eviction, deportation, compulsory evacuation, compulsory transfer of domicile, deportation to camps, transfer from camp to camp (not within camps).

This trauma deals solely with the removal and transit. It does not cover the means of transportation, nor the destination. Concentration camps, police stations, commissariats, "Germany," "Siberia," or any imposed destination, are coded T8; ghettoes, T5. Portions of journeys interruptod by brief stop-overs (for hours or overnight) are not coded separately, although the places of stop-overs are coded if they are places of constraint, viz. "From home we were taken to the commissariat. There we were registered and sent to (camp) Drancy." *Taken from home, and sent,* T1; commissariat, T8; from there sent (no coding) to Drancy, T8.

2. No adequate substitutes for removed stimuli. (1Ph;1Exp)

*Interpretative Expansion—(a) Enivronmental come-down.* Transfer to significantly substandard housing or improvised, inadequate shelters outside the camps (not brief stop-overs). Particulars of housing conditions outside or within camps are coded when coverable by other *T*'s.

3. Introduction of new stimuli especially in the form of human beings unpredictable in behavior and not restricted in their behavior by law, tradition or threat of complaint. (2Ph;9Exp)

*Interpretative Expansion*—Lack of recourse to law. Lack of information as to rules which were to govern the individual's present or future conduct and mode of living. *New human milieu.* (*a*) Heterogeneous masses of people, strange to *each other,* differing in age, sex, nationality, country of origin, language, social status, even if all are protagonists of the same misfortune. (*b*) Lack of recourse to law.—Follow other seven expansion items.

4. Break-up of the family group. (1A;1Exp)

5. Exclusion from the original social group, being forced into the position of out-group. (1A;2Exp)

6. Confiscation of personal property, money. (1Ph;6Exp)

### II. CULTURAL-AFFECTIVE

7. Death of relatives or ignorance of their fate. (2A;2Exp)

8. Blocking or restriction of responses to inherent or acquired drives. ("Conative traumatization"). (1Ph;7Exp)

*Interpretative expansion*—Forbiddance or restriction of responses to inherent or acquired drives (conative traumatization). States of constraint such as (*a*) arrest, (*b*) imprisonment, (*c*) "actions," "selections," raids. (*d*) Any concentration camp, prison, *or* region of settlement under constraint (Siberia, Germany for the French, etc.). Do not score the region when place of constraint is specified as concentration camp or prison. Follow for more expansion items.

9. Creation of prolonged (protracted) states of terror. (3A;10Ex)

*Interpretative expansion*—States of terror, fear, or anxiety. (*a*) Information and rumors about calamities inflicted upon other communities. See notes 1, 2, 3 below. States of (*b*) flight, (*c*) hiding and (*d*) illegal existence (false papers or assumed status as aryan, etc.) or travel. (*e*) Recourse to the black market or infringement of law or discipline under pressure of elementary needs (in medical cases code T18 only). (*f*) Protracted black-out of information through press and radio. Consummated suicide or self-inflicted injuries for reasons other than attempted suicide. (See also T15). (*g*) Hysterical screaming. Fainting. (*h*) Irrational stuporous actions (smashing things, etc.). (*i*) Mass weeping and wailing (beyond a family group) in public places. (*j*) Threats or conjectures of impending traumatization. Status of threatening danger such as being *assigned* for deportation, being under sentence, even if subsequently avoided. Being under personal threatening disfavor ("He would have killed me."). Notes: (1) If contents of communication are given in terms of accomplished codable facts with specific target—code as such and use also T9 if anxiety or fears are specifically verbalized. (2) Contents of threats, fears, or rumors are coded as T9 only. ("If the men should not show up they will take the women.") (3) *Contents* of conjectural rumors or traumatic predictions emanating from "unauthorized" persons (minor officials) are not coded but the *facts of existence of such rumors or prediction as such* is coded T9.

10. Abolition of traditions of decency and dignity by suspending the separation between the sexes and privacy for bodily care and processes of bowel movements. (4A;3Ex)

11. Depersonalization. (1A;5Ex)

*Interpretative expansion*—Ignominious treatment of the individual with the utmost disregard for his rights, standards, and values. (*a*) Making people wear ludicrous attire, (*b*) adding a uniform middle name (Sarah for women, Israel for men). (Not included: yellow star—T5; tatooing—T40). (*c*) Descriptions: "Like cattle" (not simply in cattle cars), "driven," "shoved," "like we were not human," (*d*) Derelict appearance, "as if I always looked like that." If specific *T* is given as cause, score both. (*e*) Complaints against deception and hypocrisy, mockery, gloating, disparaging qualifications and opinions, if not codable as T16.

12. Abolition of religious worship. (1A;4Ex)

13. Abolition of funeral rites or any vestige of dignity in disposal of the dead, desecration of cemeteries and utilitarian processing of human bodies. (2A;4Ex)

14. Complete blocking of habits of writing and reading. (1A;2Ex)

15. Official prohibition of any expression of hope. (1A;3Ex)

*Interpretative expansion*—(*a*) Signs or posters: "From here one does not return." (*b*) Anomy—resignation to a fate of depredation, general despair, extreme apathy. (*c*) Unsuccessful attempts at suicide. Any injuries or illness as consequence of suicidal attempts, code as T18.

16. Perfidy and betrayal. (8A;4Exp)

*Interpretative expansion*—(*a*) Compelled (accomplished) participation of victims in treacherous acts directed against members of their own families or clan. (*b*) Delivery to the enemy of people who have been previously, explicitly or in fact, given asylum, or (*c*) who were considered fellow-citizens (only if this is specified or alluded to by the interviewee as an act of treachery and betrayal). Perfidious acts of turncoats and collaborationists against "their own." If means of *inducement* ("compelling") involve sufferings covered by other *T*'s such as torturous beating, starving, or withholding of food or water, or interference with sleep, code separately in addition to T16. The status and routine of trustys and camp police are not necessarily qualified as acts of betrayal. See also 11. *Gravity score*: T16 is added to any *T* if by direct statement or by clear implication a child or children under 12 years of age are included in the target. Score as such the expressions "little children," "mothers and children," "children in arms," "women and children," "children" alone, when the age under 12 is clearly definable from the context. In cases where the word children is used simply to indicate family relationship or descendancy without definitely specified age, no gravity score should be given. But for "children" of non-specified age, such as "I and my children," "children of Jewish parents," when age below 12 is not clearly revealed—no gravity score is added.

### III. MEDICAL

17. Roll calls (appells) lasting for hours, often in rain, snow, and lasting presumably until everyone was accounted for or with the obvious and often announced purpose of increasing the mortality in the camp through respiratory ailments. (8Ph;4Exp)

*Interpretative expansion*—(*a*) Roll calls (appells) lasting for hours, often in rain or snow, or scorching heat, or forced uncomfortable postures, lasting presumably until everyone was accounted for, often with obvious disregard for prisoners' health or endurance. (*b*) Population checks in city squares under similar conditions. (*c*) Withholding of warm clothing in order to stimulate higher motility at work in response to cold. (In this case do not score T37.) Subsequently added: (*d*) Grave industrial accidents connected with compulsory labor.

18. Absence of sympathy for the sick and open threat of death to the physically exhausted and ill. (4Ph;3Exp)

19. Failure to remove the contagiously ill from the barracks, and crowding of people not yet affected into barracks inhabited by the contagiously ill. (4Ph;3Exp)

20. Forcing of patients in hospitals to share their beds with one or two other patients, most frequently of different medical classifications. (2Ph;1Exp)

*Interpretative expansion*—In hospitals or general barracks.

21. Compelling "hospital" patients to remain lying for hours with bedmates who had died during the night until a block trusty had a chance to collect for his own benefit the day's food rations of the deceased as if he were still alive. (4A;0Exp)

22. Surgical operations without anaesthetics or antiseptics. (16Ph;1Exp)

23. Performance of unnecessary operations on patients who were put under anaesthetic for a different purpose. (4Ph;1Exp)

24. "Gynecological" and rectal examinations in search of jewelry or objects of precious metal. (1Ph;0Exp)

25. Threats with X-Ray examinations to detect precious stones, possibly swallowed. (1A;1Exp)

26. Transitory deterioration of sex drive in both sexes and transitory cessation of menstruation in women. Mass sterilization of women. (8Ph;4Exp)

27. General absence of dental care. (1Ph;0Exp)

## IV. LABOR

28. Requiring forced or "slave" labor. (1Ph;2Exp)

29. Chronic overtaxing of physical resources by overwork, bad working conditions. (4Ph;1Exp)

*Interpretative expansion*—Excessive or "impossible" production quotas or exorbitant deliveries of farm products. If this theme is given together with one or more of its consequences, these consequences are additionally coded only if they fit in any of the other traumata. Does not exclude coding wih T28. Repeated coding: While T28 is coded once for each camp, T29 may be coded as many times as overtaxing of physical resources is alluded to in connection with reports of various labor tasks.

30. Forcing prisoners to steal from the factories the necessary tools and materials in order to keep the barracks in repair, while stealing of property from defense plants was classed as a major offense. (1A;1Exp)

## V. DIRECT BODILY VIOLENCE

31. Active interference with sleep and rest. (3Ph;0Exp)

*Subsequently added*: Extreme weather conditions such as heat, cold, rain, or snow unless the event is codable as T17, T37, T44.

32. Corporal punishment for personal offenses and group punishment whether the alleged offender could be found or not. (4Ph;2Exp)

*Interpretative expansion*—(*a*) Punishment administered by "authority" or "on the spot" whether for offenses of a single or several members of the group. (*b*) Confinement to dungeons. Note: This *T* is used only when the *punitive* character of such corporal violence (other than killing) is specified, otherwise it is rated as indiscriminate bodily violence, etc., and coded T33.

33. Beating and killing of randomly picked individuals, at times in order to demonstrate the fate which awaits offenders against the rules of the camp. (32Ph;7Exp)

34. Forced or inadvertent attendance at executions and floggings. (2A;2Exp)

35. "Punitive work" and "punitive sports." (1Ph;2Exp)

36. Constriction of physical space necessary for actual body movements, as a consequence of crowding. (2Ph;1Exp)

### VI. APPEARANCE, CLEANLINESS, AND DRESS

37. Insufficient clothing, clothing that did not fit. (1Ph;2Exp)

38. Failure to provide facilities for keeping clean, lack of soap and water. (1Ph;0Exp)

39. Maintenance of conditions which made it impossible for prisoners to free themselves from lice and vermin. (3Ph;1Exp)

40. Brutal shaving, bathing, and delousing processes. Tattooing of prisoners. (3Ph;2Exp)
*Interpretative expansion*—(*a*) Brutal shaving, etc. (*b*) Tattooing. Note: see also T38.

### VII. TRANSPORTATION

41. Long marches on foot after weeks and months of starvation. (8Ph;3Exp)

42. Traveling for days in overcrowded boxcars wtihout facilities or room to sit down, wash, or lie down. (3Ph;1Exp)

43. No toilet facilities in locked cars where men, women, and children were locked in together. (4A;1Exp)

44. Impossibility of removing the dead from the crowded boxcars for days en route. (3Ph;1Exp)
*Interpretative expansion*—Accumulation of dead bodies en route beyond the "nearest stop."

### VIII. FOOD

45. Compulsory change of nutritional habits both in kind of allotted food and in extreme reduction of its nutritional value and bulk. Creation of prolonged states of semi-starvation and thirst. (3Ph;1Exp)

46. Failure to provide any eating facilities such as refectories, tables, or eating utensils. Not allowing time to eat. (2A; 0Exp)

## I. APPENDIX II: NOTES ON CONTENT ANALYSIS

In addition to the procedures given in the text and in the "interpretative expansions" of the Traumatic Inventory itself, the coders adhered to the following rules:

1. All instances of traumatization fitting in any of the themes or sub-

themes of the Traumatic Inventory are coded, and included in the frequency scores.

2. No self-same event should be coded twice.

3. If a trauma statement is covered by two or more sub-themes of the same theme, it is also coded two or more times, viz. "My father became ill and died from the hard work in Dachau." Coding: Hard work of "other person"—29M; illness caused by trauma—18M; death from such illness—18M; Dachau—8M. Here illness as a consequence of a trauma and death as a consequence of such illness are covered by two different aspects of T18 and are coded separately.

4. *Prestatements* are coded, viz. "All able bodied persons were forcefully evacuated to Germany. I and my brother were sent to Buchenwald." Code prestatement: *"All able bodied persons were sent . . ."*—2M; Germany—8M. Here the prestatement by implication covers the interviewee and his brother, but it also covers many other people; therefore, the instance of his and his brother's deportation are coded *in addition* to the prestatement: "I was sent . . ." (2P); my brother was sent (2M); Buchenwald—coded once, 8P.

5. *We* and *us*—if the "other" targets are not specified, code only the incident as P (personal), viz. "We were sent to Dachau": We were sent—1P; Dachau—8P. If "other persons" are specified in addition to the narrator: "I and my whole family, my brothers and their wives and little children were loaded into busses." Code "I . . . etc."—1P. All specified "other" people are given one collective score—1M; for little children—code *gravity score,* 16M.

It appears that a team of coders devoted to content analysis of a series of interviews will in addition to the more or less stabilized instructions develop a number of rules of thumb which if consistently applied are bound to increase the reliability of the coding output of the group. Such is the case in most clinics with the scoring of protocols obtained from projective techniques and other personality tests. A mimeographed booklet containing the full coding instructions together with the Traumatic Inventory and possibly one coded interview for purposes of illustration is in preparation and will be obtainable upon request from the author.

## REFERENCES

1. ALLPORT, G. W. The Use of Personal Documents in Psychological Science. New York: Social Science Research Council, Bulletin 49, 1942. Pp. 210.
2. ALLPORT, G. W., & VERNON, P. E. A Study of Values. Boston: Houghton-Mifflin, 1931.

3. BALDWIN, A. L. The statistical analysis of the structure of a single personality. *Psychol. Bull.,* 1940, **37**, 518-519.

4. BERELSON, B. Content Analysis in Communication Research. Glencoe, Ill.: Free Press, 1952. Pp. 220.

5. BLUMER, H. Critiques of Research in the Social Sciences: I. An Appraisal of Thomas and Znaniecki's "The Polish Peasant." New York: Social Science Research Council, Bulletin 44, 1939. Pp. 210.

6. BODER, D. P. La psychologia experimental como microsociologia. *Sursum,* 1924, **1**, (Published in Guadalajara, Mexico.)

7. ———. The Influence of Concomitant Activity and Fatigue Upon Certain Forms of Reciprocal Hand Movement and Its Fundamental Components. Baltimore, Md.: Johns Hopkins Press, 1935. Pp. 121. (*Comp. Psychol. Monog.,* 1935, **11**, No. 4.)

8. ———. I Did Not Interview the Dead. Urbana, Ill.: Univ. Illinois Press, 1949. Pp. 220.

9. ———. Topical Autobiographies of Displaced People Recorded Verbatim in Displaced Persons Camps. Series of 1950, published in microcard and microfilm form. Chicago: 1951. (22 cards, obtainable from the author.)

10. ———. Topical Autobiographies of Displaced People Recorded Verbatim in Displaced Persons Camps. Series of 1953, published in microcard and microfilm form. Los Angeles: 1953. (12 cards, obtainable from the author.)

11. BROWN, A. S. Analysis of Topical Autobiographies of Displaced Persons of Christian Faith. Unpublished M.A. thesis, Illinois Institute of Technology, 1952.

12. CARTWRIGHT, D., & FRENCH, J. R. P., JR. The reliability of life-history studies. *Charac. & Personal.,* 1939, **8**, 110-119.

13. CHILD, I. L., POTTER, E. H., & LEVINE, E. M. Children's textbooks and personality development: an exploration in the social psychology of education. *Psychol. Monog.,* 1946, **60**. Pp. 54.

14. DOLLARD, J. Criteria for the Life History. New Haven: Yale Univ. Press, 1935. Pp. 288.

15. DOLLARD, J., & MOWRER, O. H. A method of measuring tension in written documents. *J. Abn. & Soc. Psychol.,* 1947, **42**, 3-32.

15a. ERON, L. D. Frequencies of themes and identifications in the stories of schizophrenic patients and non-hospitalized college students. *J. Consult. Psychol.,* 1948, **12**, 387-395.

15b. ———. A normative study of the Thematic Apperception Test. *Psychol. Monog.,* 1950, **64**, No. 9. Pp. 48.

15c. ERON, L. D., TERRY, D., & CALLAHAN, R. The use of rating scales for emotional tone of *TAT* stories. *J. Consult. Psychol.,* 1950, **14**, 473-478.

16. FEARING, F. Toward a psychological theory of human communication. *J. Personal.,* 1953, **22**, 71-88.

17. HAMMOND, P. The Psychological Impact of Unprecedented Social Catastrophe: An Analysis of Three Topical Autobiographies of Young Displaced Persons. Unpublished M.A. thesis, Illinois Institute of Technology, 1951.

18. LASSWELL, H. D., LEITES, N., *et al.* Language of Politics; Studies in Quantitative Semantics. New York: Stewart, 1949. Pp. 398.

19. LAZARSFELD, P. F. The controversy over detailed interviews—an offer for negotiation. *Pub. Opin. Quart.,* 1944, **8**, 38-40.

20. MANZER, C. W. The effect of fatigue upon variability of output of muscular work. *J. Exper. Psychol.,* 1934, **17**, 257-269.

21. MERTON, R. K., & LAZARSFELD, P. F. (*Eds.*) Continuities in Social Research; Studies in the Scope and Method of *The American Soldier*. Glencoe, Ill.: Free Press, 1950. Pp. 255.

22. MOWRER, O. H. Psychotherapy, Theory and Research. New York: Ronald Press, 1953. Pp. 700.

23. MURPHY, G. In the Minds of Men. New York: Basic Books, 1953. Pp. 306.

24. MURRAY, H. A. Explorations in Personality. New York: Oxford Press, 1938. Pp. 761.

25. PARSONS, T., & SHILS, E. A., *et al.* Toward a General Theory of Action. Cambridge, Mass.: Harvard Univ. Press, 1951. Pp. 506.

26. RECKLESS, W. C., & SELLING, L. S. A sociological and psychiatric interview compared. *Amer. J. Orthopsychiat.*, 1937, **7**, 532-539.

27. RUESCH, J., & BATESON, G. Communication, the Social Matrix of Psychiatry. New York: Norton, 1951. Pp. 314.

28. SELLING, L. S. The autobiography as a psychiatric technique. *Amer. J. Orthopsychiat.*, 1932, **2**, 162-171.

29. SPIEGELMAN, M., TERWILLIGER, C., & FEARING, F. The reliability of agreement in content analysis. *J. Soc. Psychol.*, 1953, **37**, 175-187.

30. STEPHENSON, W. A statistical approach to typology: the study of trait-universes. *J. Clin. Psychol.*, 1950, **6**, 26-38.

31. Studies in Social Psychology in World War II. Princeton: Princeton Univ. Press, 1949.

32. THOMAS, W. I., & ZNANIECKI, F. The Polish Peasant in Europe and America. New York: Knopf, 1927.

33. UHER, A. An Analysis of Four Topical Autobiographies of Mature Displaced Persons. Unpublished M.A. thesis, Illinois Institute of Technology, 1952.

34. U. S. EMPLOYMENT SERVICE. Dictionary of Occupational Titles. Washington, D. C.: U. S. Gov't Print. Off., 1939-44.

35. WARREN, R. P. The Rime of the Ancient Mariner, by Samuel Taylor Coleridge "A poem of pure imagination: an experiment in reading, by Robert Penn Warren." New York: Reynal & Hitchcock, 1946. Pp. 148.

35a. WHITE, R. K. Value analysis: A quantitative method for describing qualitative data. *J. Soc. Psychol.*, 1944, **19**, 351-358.

35b. ———. Black Boy: A value analysis. *J. Abn. & Soc. Psychol.*, 1947, **42**, 440-461.

36. WUNDT, W. Grundriss der Psychologie. Leipzig: Kröner, 1918. Pp. 414.

*Department of Psychology*
*University of California*
*Los Angeles 24, California*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 51-65.

# STUDIES IN PHYSIOGNOMIC PERCEPTION: I. EFFECT OF CONFIGURATIONAL DYNAMICS AND MEANING-INDUCED SETS ON THE POSITION OF THE APPARENT MEDIAN PLANE*[1]

*Department of Psychology, Clark University*

HEINZ WERNER AND SEYMOUR WAPNER

## A. THE PROBLEM

Nineteenth and early twentieth century psychology has treated the problem of perception mainly in terms of its structural aspects, such as intensity, extensity, form, etc. Though gestalt psychology has revolutionized our approach to perceptual problems by stressing the priority of the whole over the elements, it has, nevertheless, in the main, reinforced structural analysis, e.g., analysis in terms of figure-ground relations, part-whole relations, autochthonous laws of organization, etc.

In contrast, only a relatively small number of psychologists have been aware of characteristics of perception not describable in terms of structural but rather dynamic or "expressive" qualities. Uexküll (7), for instance, in his analysis of the world of animals and humans has pointed to functional signs of objects ("sitting-tone" of a chair); similarly, Lewin (2) refers to the so-called demand-character of things; and Werner (9) stressed the importance of dynamic aspects of perception by systematically pointing to their importance for ontogenetic psychology, language behavior and psychopathology. Because human physiognomy is the outstanding paradigm of an object universally perceived as an expressive entity, Werner proposed the term "physiognomic perception" for this general mode of cognition.

To illustrate the difference between structural and dynamic properties of a perceptual object we may refer here to the area of reversible figures. Take, for instance, the reversible figure shown in Figure 1. Gestalt psychology has treated such figure reversal descriptively by pointing to the exchange of figure and ground, the belongingness of contour to one part of a field versus another part, etc. What cannot be stated by such formal descriptions is that in one case we are perceiving a face looking toward the left and in the

---

*Received in the Editorial Office on February 3, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a research grant from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

other a face looking toward the right. In other words, over and above formal changes stressed by gestalt psychology there are differences best described as those of visual dynamics; i.e., one may state that one face has a directional quality to the left, and the other a directional quality to the right.

Such descriptions of visual dynamics remain at the purely phenomenological level as long as no provision is made to demonstrate the dynamic character of objects by objective behavioral measurement.

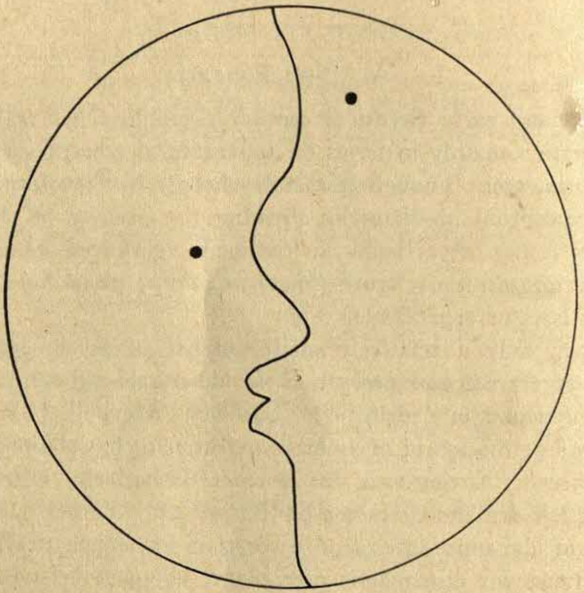


FIGURE 1
AMBIGUOUS FORM: ADAPTED FROM SCHAFER AND MURPHY (5)

It is the purpose of the following two experiments to demonstrate that "visual directional dynamics" exist as an organismic, behaviorally measurable event. The general methodology employed here is that of demonstrating differential effects of the visual dynamics inherent in figures upon their location in space, in particular their position with respect to the apparent median plane (straight-ahead).

Before going on to the particular methodology used for measuring "directional dynamics" it is necessary to refer to previous experiments concerning the apparent median plane. In these experiments it has been demonstrated that the physical location of the apparent median plane is significantly affected by the placement of a test object (8, 11). For example, if a luminescent square is placed so that its left edge is in the objective median plane and fixated, the square is perceived as displaced to the left; this means, that the straight-ahead, originally at the objective median plane, is now to the right. This perceptual displacement of the apparent median plane was measured by moving the square to the right so that the fixated edge appeared to *S* to be "straight-ahead." Similarly when the right edge is fixated the apparent median plane shifts to the left. In general, then, the apparent median plane shifts toward the center of a figure which is placed asymmetrically with respect to the objective median plane.

Further, it has been found that the magnitude of shift of the apparent median plane depends on the magnitude of asymmetrical extent: the greater the asymmetrical extent, the greater the shift of the apparent median plane toward the center of the asymmetrically placed figure (1). Thus, the efficacy of a figure to induce a shift of the apparent median plane can be described vectorially, i.e., in terms of direction and magnitude: e.g., a figure with its left edge in the objective median plane, extends to the right and shifts the apparent median plane to the right; the magnitude of shift depends on the magnitude of asymmetrical extent.

Turning now to figures with inherent dynamic properties, we assume that the presence of these properties can be demonstrated by their effect on the location of the apparent median plane. This efficacy of directional dynamics on location in space was studied in two experiments; in the first experiment directional dynamics was varied simultaneously with asymmetrical extent, in the other, directional dynamics was investigated while keeping asymmetrical extent constant and varying dynamics alone.

Both experiments had the following apparatus and procedure in common. The tests were performed in a dark room. *S* sat erect throughout all trials in a chair with side supports; his head was held in position by an adjustable head rest. In this way, position of head and trunk were kept constant throughout the trials. In both experiments luminous test figures were used. The *S*'s task, in general, was to fixate part of the test figure and to instruct *E* to move the figure so that the fixated part appeared straight-ahead to *S*. The figure was placed at a specified distance from *S* and could be moved on a track in small steps horizontally in the fronto-parallel plane. The position of the apparent median plane was measured in terms of deviation from the objective median plane. In turn, the objective median plane was determined by projecting the median sagittal plane of *S*'s head and body upon the fronto-parallel plane. Arbitrarily, positions to the left of the objective median plane were designated by minus values; positions to the right of the objective median plane were designated by plus values.

## B. EXPERIMENT 1: EFFECT OF HORIZONTAL DIRECTIONAL DYNAMICS OF FIGURES ON THE POSITION OF THE APPARENT MEDIAN PLANE[2]

In this experiment figures were used which have a directional dynamics to the left or the right. Such figures are expected to be perceived as displaced in the direction of the dynamics; the apparent median plane (straight-ahead) is expected to shift opposite the direction of dynamics. Thus, a figure having directional dynamics to left will be perceived shifted to left, or, in terms of the apparent median plane, such a figure should have the effect of shifting the apparent median plane to the right.

Specifically then, the general hypothesis of this experiment can be stated as follows: Figures having horizontal directional dynamics toward the left (toward the right) will shift the apparent median plane farther toward the right (left) than figures which do not have horizontal directional dynamics, such as a square.

### 1. *Procedure*

Five luminescent test figures, four with inherent directional dynamics and one lacking it, were used. The figures were placed two meters from *S*. The figures, 20 cm. high and 20 cm. wide, were: an equilateral triangle with apex pointing right; a right profile of a face; a square; an equilateral triangle with apex pointing left; a left profile of a face. Each figure was used in two ways: once the extreme right part of the figure (edge, point) was placed in the objective median plane and fixated; once the extreme left part of the figure was placed in the objective median plane and fixated. There were thus 10 conditions which are pictured in Table 1. This table also describes the factors introduced into the particular experimental condition, the vectorial direction of the factors operating, and states the predicted relative rank position of the apparent median plane for left extent and for right extent separately.

Certain predictions can be made from the direction of the two vectors assumed to be operating as to the rank position of the apparent median plane. If one were to make predictions for all 10 conditions one would have to make assumptions not only as to direction but also as to magnitude of these vectors. Since we make no assumption as to the relative potencies of the two vectors operating, we confine our prediction to the relative position of the apparent median plane within the conditions of left extent (Conditions 1-5) and within the conditions of right extent (Conditions 6-10). First of all, asymmetrical extent operates to shift the location of the apparent

---

[2] The experiment was carried out by Jan H. Bruell and Alvin G. Goldstein.

TABLE 1

EXPERIMENT 1: TEST CONDITIONS, FACTORS OPERATING, AND PREDICTED RELATIVE POSITION OF APPARENT MEDIAN PLANE

| Test Condition | Figure and Placement OMP* | Factors Operating | | Vectors Operating** | | Predicted rank*** position of apparent median plane from left (1) to right (3) | |
|---|---|---|---|---|---|---|---|
| | | Asymmetrical Extent | Directional Dynamics | Asymmetrical Extent | Directional Dynamics | Left Extent | Right Extent |
| 1 | | Left | Right | ← | ← | 1 | - |
| 2 | | Left | Right | ← | ← | | |
| 3 | | Left | None | ← | None | 2 | - |
| 4 | | Left | Left | ← | → | 3 | - |
| 5 | | Left | Left | ← | → | | |
| 6 | | Right | Right | → | ← | - | 1 |
| 7 | | Right | Right | → | ← | | |
| 8 | | Right | None | → | None | - | 2 |
| 9 | | Right | Left | → | → | - | 3 |
| 10 | | Right | Left | → | → | | |

* Objective median plane

** Vectors are assigned in terms of predicted direction of shift of apparent median plane only.

*** No assumptions are made as to differential potency of the directional dynamic vectors.

median plane toward the center of the test figure. Furthermore, with respect to directional dynamics, our expectation is that there should be a shift of the apparent median plane in a direction opposite to that of the dynamics. Accordingly, with test Conditions 1-5 (left extent), the expectation is that the apparent median plane should fall most to the left for Conditions 1 and 2, most to the right for Conditions 4 and 5, and between these extremes for Condition 3. No assumption is made as to the differential efficacy of face versus triangle. Analogously, the predicted rank order for test Conditions 6-10 (right extent) is (from left to right): 6 and 7; 8; 9 and 10.

Forty *S*s were tested under each of the 10 experimental conditions. One trial was given for each test condition. A 10 x 10 replicated latin-square design was used. Four *S*s, two men and two women, were tested in each of the 10 sequences of conditions.

## 2. *Results*

A summary of the analysis of variance is presented in Table 2a. It can be seen that the *F* test for conditions is significant beyond the .01 level.

Table 2b presents the means, mean differences, and tests of significance. The results are in accordance with the predictions: for asymmetrical extent left the apparent median plane is relatively most to the left for Conditions 1 and 2 (—6.4 cm., —6.1 cm.) next for Condition 3 (—5.9 cm.), and farthest to the right for Conditions 4 and 5 (—4.1 cm., —3.8 cm.); for asymmetrical extent right the apparent median plane is relatively most to the left for Conditions 6 and 7 (+.2 cm., +1.8 cm.), next for Condition 8 (+3.7 cm.), and farthest to the right for Conditions 9 and 10 (+4.4 cm., +4.6 cm.).

TABLE 2a

EXPERIMENT 1: EFFECT OF DIRECTIONAL DYNAMICS AND ASYMMETRICAL EXTENT ON POSITION OF APPARENT MEDIAN PLANE

| | | F tests | | |
|---|---|---|---|---|
| Source of variation | *df* | Mean Square | *F* | *P* |
| Individuals | 39 | 269.23 | —— | —— |
| a. Sex | 1 | 58.98 | <1.00* | >.05 |
| b. Sequence | 9 | 228.90 | <1.00* | >.05 |
| c. Sex and Sequence | 9 | 377.27 | <1.00* | >.05 |
| d. Individuals within sex and sequence | 20 | 419.04 | 16.87** | <.01 |
| Conditions | 9 | 837.19 | 33.70** | <.01 |
| Ordinality | 9 | 41.10 | 1.65** | >.05 |
| Square Uniqueness | 72 | 16.36 | <1.00** | >.05 |
| Error | 270 | 24.84 | | |
| Total | 399 | | | |

*Tested against individuals within sex and sequence.
**Tested against error.

Within the conditions of left asymmetrical extent (Conditions 1-5) the figures with left directional dynamics (Conditions 1 and 2) differ significantly from those with right directional dynamics (Conditions 4 and 5); however, only Conditions 4 and 5 differ significantly from Condition 3 (no dynamics) ($p < .05$, one-tailed test). A similar picture obtains within the conditions of right asymmetrical extent (Conditions 6-10): the figures with left directional dynamics (Conditions 6 and 7) differ significantly from those with right directional dynamics (Conditions 9 and 10); however, only Conditions 6 and 7 differ significantly from Condition 8 (no dynamics) ($p < .05$; one-tailed test). In general, it is quite clear that, as predicted, the apparent median plane shifts relatively to the left for figures with "right directional dynamics," and relatively to the right for figures with "left directional dynamics."

The data also provide confirmatory evidence that the apparent median plane shifts toward the center of an asymmetrically placed figure. All 25 comparisons between left and right asymmetrical extent are significant beyond the .01 level of confidence.

TABLE 2b

EXPERIMENT 1: EFFECT OF DIRECTIONAL DYNAMICS AND ASYMMETRICAL EXTENT ON POSITION OF APPARENT MEDIAN PLANE

Means, Mean differences, *t* tests*

| | Test condition<br>Asymmetrical extent | Figure | Mean position of apparent median plane (cm.)** | Mean difference<br>1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Left | Triangle Right† | —6.4 | — | .3 | .5 | 2.3 | 2.6 | 6.6 | 8.2 | 10.2 | 10.8 | 11.0 |
| 2 | | Face Right | —6.1 | | — | .2 | 2.0 | 2.3 | 6.3 | 7.9 | 9.9 | 10.5 | 10.7 |
| 3 | | Square | —5.9 | | | — | 1.8 | 2.1 | 6.1 | 7.7 | 9.7 | 10.3 | 10.5 |
| 4 | | Face Left | —4.1 | | | | — | .3 | 4.3 | 5.9 | 7.9 | 8.5 | 8.7 |
| 5 | | Triangle Left | —3.8 | | | | | — | 4.0 | 5.6 | 7.6 | 8.2 | 8.4 |
| 6 | Right | Triangle Right | + .2 | | | | | | — | 1.6 | 3.6 | 4.2 | 4.4 |
| 7 | | Face Right | +1.8 | | | | | | | — | 2.0 | 2.6 | 2.8 |
| 8 | | Square | +3.8 | | | | | | | | — | .6 | .8 |
| 9 | | Face Left | +4.4 | | | | | | | | | — | .2 |
| 10 | | Triangle Left | +4.6 | | | | | | | | | | — |

*The standard error of the difference used in computing *t* was derived from the error variance (*df* 270) indicated in the F table (Table 2a). Mean differences required for significance (one-tailed test): 1.8 ($p = .05$); 2.2 ($p = .01$).

**Positions to the left of the objective median plane are designated by "—" values; positions to the right of the objective median plane are designated by "+" values.

†Triangle pointing to the right.

Further, from the results the inference can be drawn that for the particular test figures used in this experimental situation, asymmetrical extent is more potent than directional dynamics; i.e., in cases where the vector of asymmetrical extent is opposite that of dynamics (Conditions 4, 5, 6, 7) the relative position of the apparent median plane is determined by extent (Conditions 4 and 5 fall to the left of Conditions 6 and 7). The question may be raised whether the greater potency of asymmetry over dynamics is a general one. Some evidence is available that this is not the case: in one experiment dynamic figures were presented under labyrinthian stimulation (the *S* and the figure rotate together around a vertical axis); it was found that here the dynamic factor was more potent than the factor of asymmetrical extent (3).

Finally, though significant differences were obtained for "directional dynamics" inherent in such figures as a triangle pointing to the left and right and a face pointing to the left and right, the results are somewhat equivocal as to interpretation. Since the "physiognomic" figures differ in shape from the square, it might be questioned whether these effects are a function of the particular shape of the figures rather than of some dynamic factor. The interpretation would be less ambiguous if it were possible to induce opposite dynamics in a single stimulus pattern and then demonstrate significantly different shifts in the apparent median plane. Since an identical stimulus pattern would be involved, differences in effect could then be unequivocally attributed to the dynamics per se. These considerations led to the methodology of the second experiment.

## C. EXPERIMENT 2: EFFECT OF HORIZONTAL DIRECTIONAL DYNAMICS INDUCED BY A SET ON THE POSITION OF THE APPARENT MEDIAN PLANE

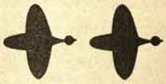
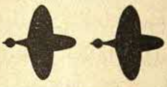
To evaluate whether shifts in apparent median plane due to directional dynamics occur independent of differences in stimulus pattern, it was necessary to use a single stimulus pattern which could be varied in directional dynamics. This crucial requirement was satisfied by the use of a stimulus pattern which, according to instructions, was to be perceived in one of two ways, i.e., with directional dynamics to the left or with directional dynamics to the right. Briefly, the instruction induced a meaning to be given to the stimulus pattern which determined the direction of the dynamics.

The hypothesis tested was as follows: opposite directional dynamics induced in a test figure by different meaning-sets have organismic effects reflected in opposite shifts of the position of the apparent median plane; when the induced directional dynamics of the pattern is to the left, the apparent median plane should shift relatively to the right in comparison with the same pattern, when the induced directional dynamics is to the right.

### d. *Procedure*

The ambiguous stimulus pattern used was seen, according to instructions, either as the silhouette of two flying-birds or of two flying-airplanes. This pattern was presented in two orientations (see Table 3): test Pattern 1—birds flying right (or airplanes flying left); and test Pattern 2—birds flying left (or airplanes flying right).

TABLE 3
EXPECTED SHIFTS OF APPARENT MEDIAN PLANE

| TEST PATTERN | DYNAMICS | | EXPECTED SHIFT OF THE APPARENT MEDIAN PLANE | |
|---|---|---|---|---|
| | BIRD-SET | AIRPLANE SET | BIRD-SET | AIRPLANE SET |
| #1  | → | ← | ← | → |
| #2  | ← | → | → | ← |

A square of black velvet paper was attached to the milk glass face of a light box and the center portion cut out so that in darkness a solid luminous silhouette of the test pattern was visible. The light box contained a 5-watt bulb operating at 30 volts and was movable on a track in the fronto-parallel plane 3.2 meters from *S*. To increase lability of perception the test patterns were presented as slowly flashing luminous silhouettes: the light interval was .5 sec.; the dark interval .5 sec. At the beginning of each trial the test pattern was placed so that the center (dark space) was in the objective median plane. At *S*'s instruction *E* moved the test pattern until the center appeared "straight-ahead" to *S*.

A total of 80 *S*s were used. Each test pattern (Nos. 1 and 2) was adjusted to the "straight-ahead" under one of two instructional sets: 40 *S*s (20 men; 20 women) were told that the test patterns were the pictures of two birds flying one behind the other (bird-set); the remaining 40 *S*s (20 men; 20 women) were told that the test patterns were the pictures of two airplanes flying one behind the other (airplane-set).

Preliminary instructions were given before *S* entered the experimental room to enforce the desired instructional set: the bird-set group was shown a picture of a flying mallard duck and then told they would see silhouettes of birds in flight; the airplane-set group was shown a model plane and told they would see silhouettes of flying airplanes.

There were two consecutive trials for each test pattern. The measure used in the analysis was the mean of two trials for each pattern. For half of the men (women) in the bird-set group and for half of the men (women) in the airplane-set group, the sequence of test patterns was 1, 2; for the other half it was 2, 1.

The hypothesis is made explicit in Table 3. According to the hypothesis, for test Pattern 1, the expectation is that under bird-set with directional dynamics to the right, the apparent median plane will shift to the left; under airplane-set with directional dynamics to the left, the apparent median plane will shift to the right. For test Pattern 2 the expectations concerning the shifts are opposite that of test Pattern 1.

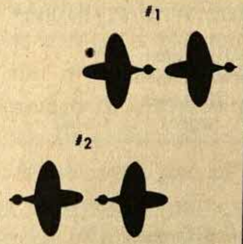
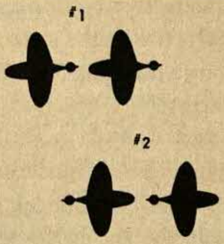
### 2. *Results*

The data were analyzed by Chi-square and by *t*-tests. For purpose of Chi-square analysis the responses of the *S*s were classified in one or the other of two categories: Category 1—test Pattern 1 adjusted to the left of test Pattern 2: Category 2—test Pattern 1 adjusted to the right of test Pattern 2. The tabulation based on these categories is presented in Table 4. Since a heterogeneity analysis (6) showed no difference between the sexes, the data were pooled. Of the *S*s in the airplane-set group, we expect the majority to adjust test Pattern 1 to the right of 2. This expectation was verified: 26 *S*s adjusted test Pattern 1 to the right of 2, and only 14 *S*s adjusted test Pattern 1 to the left of 2. In contrast, of the *S*s in the bird-set group, we expect the majority to adjust test Pattern 1 to the left of 2; actually 25 *S*s in this group adjusted test Pattern 1 to the left of 2 and only 15 adjusted 1 to the right of 2. The Chi-square was significant ($p < .01$; one-tailed test) providing support for the hypothesis.

The *t*-test analysis was based on a "difference score" obtained by subtracting the mean position of the apparent median plane under conditions of test Pattern 2 from the mean position of the apparent median plane under conditions of test Pattern 1 (1 minus 2). The sign of the expected "difference score" for bird-set compared with airplane-set can be found in Table 5. For each *S* such a difference score was computed, and the significance of the difference between the mean "difference score" for the bird-set group and

TABLE 4

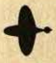
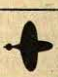
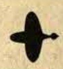
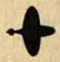
DIRECTIONAL DYNAMICS AND SET

| | RELATIVE PLACEMENT OF TEST PATTERNS | |
|---|---|---|
| | #1 TO RIGHT OF #2 | #1 TO LEFT OF #2 |
| |  |  |
| UNDER AIRPLANE-SET (N=40) | 26 | 14 |
| UNDER BIRD - SET (N =40) | 15 | 25 |

$\chi^2 = 6.05$ $p < .01$ (one-tailed test)

TABLE 5

EXPECTED SIGN OF DIFFERENCE SCORES

| SET | TEST PATTERN (SCHEMATIC) | VECTORS: EXPECTED SHIFT OF APPARENT MEDIAN PLANE | SIGN OF EXPECTED DIFFERENCE SCORES (#1-#2) |
|---|---|---|---|
| BIRD | #1  | (relatively left: -) | − |
| | #2  | (relatively right: +) | |
| AIRPLANE | #1  | (relatively right: +) | + |
| | #2  | (relatively left: -) | |

the airplane-set group was evaluated by a *t*-test. The results, presented in Table 6, are in the expected direction, —2.0 cm. for bird-set and +2.5 cm. for airplane-set;[3] the difference between these means is significant ($p < .05$; one-tailed test).

TABLE 6

SIGNIFICANCE OF DIFFERENCE BETWEEN MEAN DIFFERENCE SCORES UNDER BIRD-SET COMPARED WITH PLANE-SET*

| | Mean difference score (No. 1 — No. 2) | Difference between mean difference scores | *df* | *t* | *P*** |
|---|---|---|---|---|---|
| Bird-set | —2.0 cm. | 4.5 cm. | 78 | 1.81 | $.05 > P > .025$ |
| Plane-set | +2.5 cm. | | | | |

*The prediction is that the mean difference (No. 1 — No. 2) for bird-set will be negative and for airplane-set will be positive; the adjustments of the apparent median plane which are to the right of the objective median plane are designated +, those which are to the left of the objective median plane are designated —.

**The probability is given for a one-tailed test since the direction of the mean differences was hypothesized.

It has to be recognized that an experiment of this kind which deals with sets induced by instruction cannot be thoroughly controlled; i.e., it cannot be ruled out that some *S*s because of previous experience and predilections, may have already acquired a set contrary to that requested by instruction. To illustrate, in a pilot study, one *S* just after being given the instructions that he would see an airplane exclaimed at the sight of the test pattern "Oh, I have just been hunting, those, to me, are ducks!" It can be argued that the strong reversals we found in a few cases are explainable not in terms of inefficacy of physiognomics, but rather in terms of dynamics in the opposite direction due to such contrary attitudes on the part of the subject.[4]

In sum, the evidence from both statistical analyses provide support for the hypothesis that opposite directional dynamics induced by different meaning sets in the same configuration have organismic effects reflected in opposite shifts of the apparent median plane: when the directional dynamics is to the right, the apparent median plane shifts to the left and vice versa.

---

[3] We may add that the mean positions of the apparent median plane for stimulus Patterns 1 and 2 respectively are +1.4 and +3.4 for bird-set, and +2.8 and +.3 for airplane-set. As one will note all positions of the apparent median plane are to the right of the objective median plane, a fact not reflected in the "difference scores" themselves. This is of interest because analogous systematic effects have been found with apparent verticality, which we have interpreted from the viewpoint of spatial asymmetries characteristic of the group. However, this whole problem of spatial asymmetries, important as it is, is in need of more comprehensive study.

[4] Since such contrary attitudes may be the basis for the small number of extreme reversals found, and since the *magnitude* of the reversal affects the *t*-test rather than the Chi-square, it is possible to argue that Chi-square, here, is the more appropriate test.

## D. Discussion

Whenever students of psychology have tried to show the existence of dynamic properties of objects they have done so mainly by phenomenological analysis. In contrast, the present study aims at a demonstration of such dynamic properties by means of objective behavioral measurement.

The methodology employed here is similar to that used in previous studies involving the position of the apparent median plane. There it had been concluded that objects simply by their position in space may cause a change in the organismic state which is reflected in shifts of the position of the apparent median plane (8, 11). In the two experiments reported here we have demonstrated that the so-called "directional dynamics" of perceived objects also affects the position of the apparent median plane: if directional dynamics is to the left, the apparent median plane shifts relatively to the right, and vice versa.

At present no more than a rather crude and tentative interpretation can be attempted. The framework for this interpretation is that of the sensory-tonic field theory of perception (10). Our general assumption is that visual dynamics affects the equilibrial state of the organism in a manner which may be best described by a hypothetical vectorial diagram shown in Figure 2. We assume that the directional dynamics exerts a "pull" on the organism in the di-

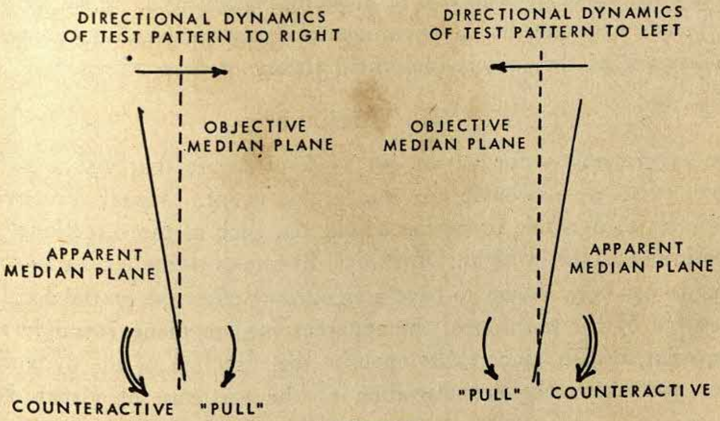


FIGURE 2
Hypothetical Schema: Shift of Apparent Median Plane Under Left and Right Directional Dynamics

rection of the dynamics, this is counteracted by an organismic pull in the opposite direction. In keeping with our previous conceptualization (see 11, p. 298), displacement of the apparent median plane is in the direction of the counteractive force.[5]

In connection with this schematization, it might be advantageous to refer to the significant analysis of H. H. Price (4) on the rôle of "muscular sensation" as a basis for the experience of forces inherent in the objective world. He contends that this experience is mediated by muscular strains and stresses.

> When I experience a push from without inwards, I also experience a counter-push from within outwards: and when I experience an outward pull (as when I am dragged with a rope) I always experience a counter-pull whose felt direction is inwards. And this two-sidedness is still present whether the initiative (so to put it) is my own or not. I may be trying to move a heavy body out of my way; or a heavy body may bump into me, without any initiative on my part. In either case, I experienced a muscular sensation. In either case this vectorial character is two-sided; it is both from without inwards and within outwards.

The only addition to this statement we would venture to make is that the assumption of the "two-sided vectorial character" as underlying the perception of dynamics is not restricted to the tactual-kinaesthetic field but applies to all sense modalities including the visual as discussed here.

The last point we wish to make refers to the problem of cognition in relation to organismic theory. We feel that this study provides—at least within the restricted area of visual perception—a methodology for studying objectively the effect of induction of meaning on percepts. This study represents, then, a first step in the experimental investigation of meaning within the framework of the sensory-tonic field theory of perception (10).

## E. Summary

Two experiments were carried out to demonstrate that visual directional dynamics exists as a behaviorally measurable event. Visual directional dynamics refer to dynamic properties of objects, such as the directional quality in a picture of a bird in flight. Directional dynamics to the left or to the right of a figure has been shown to have a significant effect on spatial localization as measured by the position of the apparent median plane (straight-ahead): the apparent median plane shifts opposite the direction of the dynamics.

Over and above the demonstration of the existence of visual dynamics

---

[5] At the moment we leave open the question of whether such an organismic pull is restricted to the ocular muscular system or involves a wider bodily distribution.

through behavioral measurement, the study shows how meaning, induced by instruction, can be treated experimentally and interpreted within an organismic theory, viz., the sensory-tonic field theory of perception.

## REFERENCES

1. **Bruell, J. H.** Visual egocentric localization. Ph.D. dissertation, Clark University, 1953.
2. **Lewin, K.** Dynamic Theory of Personality. New York: McGraw-Hill, 1935.
3. **Morant, R. B.** Factors influencing the apparent median plane under labyrinthian stimulation. Ph.D. dissertation, Clark University, 1952.
4. **Price, H. H.** Touch and organic sensation. *Proc. Aristotelian Soc.*, 1944, N. S. **44**, 1-30.
5. **Schafer, R., & Murphy, G.** The rôle of autism in a visual figure-ground relationship. *J. Exp. Psychol.*, 1943, **32**, 335-343.
6. **Snedecor, G. W.** Statistical Methods. Ames, Iowa: Collegiate Press, 1946.
7. **Uexküll, J. von.** Theoretische Biologie. Berlin: Springer, 1928.
8. **Wapner, S., Werner, H., Bruell, J. H., & Goldstein, A. G.** Experiments on sensory-tonic field theory of perception: VII. Effect of asymmetrical extent and starting positions of figures on the visual apparent median plane. *J. Exp. Psychol.*, 1953, **46**, 300-307.
9. **Werner, H.** Comparative Psychology of Mental Development. Chicago: Follet, 1948.
10. **Werner, H., & Wapner, S.** Toward a general theory of perception. *Psychol. Rev.*, 1952, **59**, 324-338.
11. **Werner, H., Wapner, S., & Bruell, J. H.** Experiments on sensory-tonic field theory of perception: VI. Effect of the position of head, eyes, and of object on the position of the apparent median plane. *J. Exp. Psychol.*, 1953, **46**, 293-299.

*Department of Psychology*
*Clark University*
*Worcester 10, Massachusetts*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 67-70.

# TESTPROXIMAL ANALYSIS, TERMINOLOGICAL ABSORPTION, AND CONCEPTUAL PURIFICATION*

*Department of Psychology, University of Lund*

---

ÅKE BJERSTEDT

---

## A. TESTPROXIMAL AND TESTDISTAL VARIABLES

The present paper deals with some problems of psychology terminology. From a test situation we get many different variables. Schematically we may distinguish between two main categories: testproximal and testdistal variables. We may also distinguish two classes of testproximal variables: primary and secondary. The immediate reactions of the subjects form the primary variables. In turn, the secondary variables may be said to be created by these immediate reactions when combined in one way or another: (*a*) in graphic representation (curves, schemas etc.), (*b*) in statistical treatment (means, coefficients of correlation etc.), or (*c*) in logical categorization. The testdistal variables, lastly, are obtained by transferring the primary and secondary variables to a more general psychological language.

Certainly it is desirable gradually to create a common scientific terminology, into which all different primary and secondary variables may be "translated." Several attempts to establish such more general systems of interpretative variables have been made. However, we often hear reports of low correlations between different tests, which have been said to measure the same psychological "trait"—e.g., "emotionality," "introversion," "neuroticism," "honesty." Among other things, such facts show the danger of too hastily "translating" the test data into a more general language and thus speak in favour of a testproximal analysis. Sometimes it seems wise to use an *explicit hypothetical interpretation*: such are the primary data, and such are the corresponding hypotheses. Although the interpretative hypothesis may appear to be an inferior one, nevertheless the *basis of the hypothesis*, i.e., the empirical fact upon which the hypothesis was founded, is left undisturbed.

## B. TERMINOLOGICAL ABSORPTION

Connected with the general principle to distinguish as clearly as possible between testproximal and testdistal variables is the more special principle to *keep the terms of the testproximal variables free from the interpretative*

---

*Received in the Editorial Office on February 3, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

*hypotheses.* In cases where the operative terms are "coloured" by the interpretations, it is difficult both to *disconnect the conventional interpretative hypothesis* and to *introduce new hypotheses.* While most Rorschach terms stand up to the demands of testproximal freedom of hypotheses (detail responses, colour responses, etc.), other test terminologies (e.g., part of the Szondi terms) are more "coloured."

Many unproductive psychological discussions are caused by a *conflict between operational definitions and a pre-operational terminology.* We have for a long time wanted to measure intelligence. Gradually we have had to admit that what we really measure with our intelligence tests is something more restricted than we originally set out to find. Although some workers therefore want to speak about "test intelligence" or "Binet intelligence," most psychologists still call the measured variable "intelligence," thereby sometimes confusing people with other ideas of intelligence. In a certain way every term is correct, as long as it is operationally defined and consistently used. On the other hand, nobody can escape the power of linguistic tradition; although operationally defined, terms taken from everyday language easily "absorb" parts of their everyday meaning. We might here speak of a *terminological absorption,* i.e., an absorption of ideas, connected with the traditional use of the term, but irrelevant to the operational definition. The obvious defense against such terminological absorption is to create new terms. It is clear that neologisms are often very impractical: we easily split up our world in disconnected, "closed" operational systems with difficulties of intercommunication. Of course our aim is linguistic unity combined with exactness. However where it is impossible to arrive at both, it does not seem unwise to prefer heterogeneous exactness to inexact unity. What is the use of an artificial unity, created by using the same vague terms for disparate operational realities? The usual result seems to be confusion and unnecessary discussion.

## C. Conceptual Purification

Operational definition means sometimes a *fictive isolation* of certain parts of a complex empirical content. From the *descriptive concept* $A = f(a, b, c, d \ldots)$, where the four written variables denote the main attributes, we obtain *constructed concepts* by an isolating choice of variables: $x = f(a)$; $y = f(b, d)$; etc. (cf. 8, p. 104). We may then speak of a "neglective fiction" (Vaihinger) or of a *"conceptual purification"* (Hans Larsson; *ibid.*). The point stressed above could then be phrased thus: *the conceptual purification ought to be followed by a terminological purification*

—among other things in order not to be misleading when the concept is confronted with empirical reality.

Sometimes we meet the fictive isolation in a special form as *antitypical or bipolar purification.* Useful as such a thought model may be, the danger is that, unskilfully applied, the purified concepts will be apprehended as or confounded with reality. Elsewhere I have shown the risk of the careless use of such bipolar concepts in socio-psychological work (1). The examples I selected were the following term-couples: (*a*) *Jennings* (3); the more personal evaluation in "psychegroups" ("psyche-tele"), *vs.* the more objective evaluation in "sociogroups" ("socio-tele"); (*b*) *Koskenniemi* (4): the more personal "Dimension der Nähe und Ferne," *vs.* the more objective "Dimension der Distanz"; (*c*) several authors (cf. 2, p. 155, and the writings of Bogardus): "horizontal distance," denoting the degree of intimacy, *vs.* vertical distance," denoting hierarchical position. Although all these conceptual polarities may be good analytical instruments, it is necessary to make clear that we deal with abstractions and fictive purifications. With reference to Jennings Moreno has warned of too sharp a conceptual cleavage and shown that we have to talk about a "psycho-socio continuum," where the personal evaluation is more prominent in one situation, less prominent in another (6). Similar warnings may be directed against careless use of the other two conceptual couples. As I have shown earlier (1, p. 71), a picture of the conceptual interrelatedness in these cases may be given with the aid of one of Alf Nyman's graphical models of problem-solving, viz., that one illustrating Lloyd's "serial mediation" (5), called by Nyman the "method of *maxima* and *minima*" or the "redoubled method of zero" (7, p. 28; cf. Figure 1). Here one limit ($I^{max}$—$S_0$) may represent the fictive case when a group formation is directed only by objective—i.e., inter-subjective—or vertically differentiating factors ($I^{max}$ = maximum *i*nter-subjective factors;

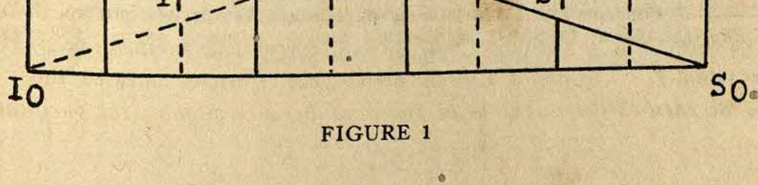


FIGURE 1

$S_0$ = minimum or zero subjective factors). The other—likewise extrapolated—limit ($S^{max}$—$I_0$) denotes the case when only personal-subjective or horizontally differentiating factors direct the process of group formation. The empirical cases are all mixed forms, where now one factor-constellation may overweigh, now the other may preponderate. It should be clear that the three term couples here mentioned are not altogether identical as to meaning. But as to their conceptual interrelations, two by two, they are quite similar. The thought model of the "serial mediation" is equally applicable to all of them as well as to a long series of other psychological bipolarities.

## D. SUMMARY

The present paper deals with some problems of terms and concepts in psychology. In connection with the general principle, as clearly as possible to distinguish between testproximal and testdistal variables, i.e., the starting-point and the result of the interpretative procedure, the more special principle to keep the terms of the testproximal variables free from the interpretative hypotheses is stressed. The conflict between operational definitions and pre-operational terminology is discussed. Irrelevant terminological absorption may be diminished by allowing a conceptual purification to be followed by a terminological purification. Lastly the dangers of careless bipolar purification are mentioned, and the conceptual interrelatedness in these cases is illustrated with a special graphical device (Nyman: "redoubled method of zero") and some empirical examples.

## REFERENCES

1. BJERSTEDT, A. Sociometric Studies of Factors Related to School Children's Choices of Work Mates. Unpublished fil. lic. thesis, University of Lund, May, 1952. (In Swedish.)
2. CUVILLIER, A. Manuel de Sociologie. Paris: Presses Univ. France, 1950.
3. JENNINGS, H. H. Sociometric differentiation of the psychegroup and the sociogroup. *Sociometry,* 1947, **10**, 71-79.
4. KOSKENNIEMI, M. Soziale Gebilde und Prozesse in der Schulklasse. Helsinki: Druckeri, 1936.
5. LLOYD, A. H. The logic of antithesis. *J. Philos., Psychol., & Sci. Meth.,* 1911, **8**, 281-289.
6. MORENO, J. L. Progress and pitfalls in sociometric theory. *Sociometry,* 1947, **10**, 260-272.
7. NYMAN, A. Problèmes et solutions en philosophie. *Theoria,* 1940, **6**, 1-42.
8. ———. Concepts-limites et "culture pure" dans l'étude des lettres. *Bull. Soc. Roy. Let. Lund,* 1950-1951, **3**, 99-119.

*Hornsgatan 3*
*Ystad, Sweden*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 71-72.

# A BRIEF NOTE ON LEARNING "PRINCIPLES" AND READING*

*Institute of Educational Research, Indiana University*

Edmund Mech[1]

It is generally recognized that the psychology of reading is a learning process and should be subject to study as such. It is well known, too, that the quantity of "work" in the foregoing area is approaching staggering proportions.

More particularly, there has been a pronounced emphasis in the reading area upon remedial procedures. The term remediation can be used interchangeably with terms such as reeducation or re-learning.

There are, for example, several general methods of teaching reading. According to Burt and Lewis (1) these methods are (*a*) Alphabetic, (*b*) Phonic, (*c*) Visual, and (*d*) Kinaesthetic.

Using each of the above mentioned techniques Burt and Lewis (1) carried out an experiment to test their relative merits in teaching reading to the retarded.

It was concluded that, on the whole, the better method for teaching retarded readers is one which is predominantly visual (the look-and-say method). The kinaesthetic by itself proved less satisfactory; the alphabetic and phonic, the least effective of all in that order.

They point out, however, that the most successful method was only 10 per cent more effective than the least successful.

The indication here that one treatment is superior to other treatments for a selected sample is in itself a nice bit of knowledge, but unfortunately does not demonstrate the mechanisms responsible. This type result appears akin to applying a "therapeutic" procedure, obtaining a "recovery" and not being able to account or explain *how* it was brought about.

---

*Received in the Editorial Office on February 12, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The following are revisions to the Mech article titled, "Factors Influencing Routine Performance under Noise: I. The Influence of 'Set'," published in the April, 1953, issue of this Journal, pp. 282-298.

The author would like to clarify that the technique described in his article was directly modelled after the technique described by Baker (1) in an earlier work. Also, on p. 283, line 15, the phrase, "According to Baker (1)," should precede the sentence "one of the reasons . . . data have been gathered." On p. 283, line 17, the phrase, "Baker (1) further points out," should precede the sentence "that the subject . . . unfortunate aspect of the experiments in this area."

However, many investigators have wrestled with the problem of explaining the reading process. For example, Buswell (2), in a monograph dealing with the improvement of adult reading, states as follows: "Reading is simply a process of interpreting a certain kind of visual experience in which conventional symbols on a printed page are substituted for common objects in the environment. The ordinary laws of perception apply as in any other visual experience" (p. 144).

This may be so, but until it is demonstrated precisely what the "ordinary laws of perception" consist of, and an attempt made to handle them in a testable manner, the concept will remain somewhat nebulous.

Witty (3) describes 10 "principles" of learning which he maintains can be generally useful in our educational efforts. These "principles" were derived by Witty from the Army literacy program during World War II. For the sake of brevity they are not listed here. However, one of Witty's principles is concerned with the development and stimulation of strong interest and motivation in the learner.

The above principle informs us "what ought to be" but does not demonstrate the specific operations necessary to realize this objective.

It is one thing to state that the organism must be motivated, and interest must be developed. What variables, however, are required to motivate?

It has often been pointed out that the understanding of how human organisms learn is a most important problem in educational method. To properly assault this problem educational psychology must begin by taking behavior as its basic datum and describing specific segments of behavior under certain conditions in relation to the operations performed upon the organism.

Only when the investigator dealing with the psychology of reading has described in terms of operations the antecedent conditions and performance changes in relation to these antecedent conditions do we have an objective relationship between environmental and response variables.

## REFERENCES

1. Burt, C., & Lewis, R. Teaching backward readers. *Brit. J. Educ. Psychol.*, 1946, **16**, 116-132.
2. Buswell, G. T. How adults read. *Suppl. Educ. Monogr.*, No. 45. Chicago: Univ. Chicago Press, 1937.
3. Witty, P. Principles of learning derived from the result of the Army's program for illiterate and non-English-speaking men. *Adult Educ. Bull.*, 1947, **11**, 131-136.

*Institute of Educational Research*
*Indiana University*
*Bloomington, Indiana*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 73-78.

# DOMINANCE IN THE PERSONALITY OF THE STUDENT NURSE AS MEASURED BY THE A-S REACTION STUDY*

*Department of Psychology, Santa Barbara College, University of California*

ALMA P. BEAVER[1]

## A. THE PROBLEM

The present study is part of a more extensive investigation into the efficiency with which a battery of standardized tests predicts success of students in a college of nursing. Findings from two previous studies (2, 3) which utilized the *MMPI* and the Kuder Preference Record in identifying personality attributes and interest preferences, respectively, have been reported. In each case the experimental group was composed of applicants to a college of nursing, matched by a control group of women in a liberal arts college majoring in education curricula. Analysis of the responses of the respective groups to the individual items furnished the basis for the discrimination in the *MMPI* rather than the conventional profile. Likewise in the Kuder Preference Record, analysis of item choices rather than total scores for the nine areas proved to be the most fruitful approach.

A group of 66 items in the *MMPI,* all significant at the .05 level of confidence or better, pointed to personal qualities on the part of the nurses which tend to the mannish, and to conventional attitudes. Twenty-three of these items were significant at the .01 level of confidence. Also substantiated in the study was the commonly accepted belief in the greater stability of the nurse as contrasted with women in general. The Kuder Preference Record study yielded a total of 96 choices made to 76 items which discriminated at the .01 level of confidence. The student nurse shows a preference for science, anything pertaining to the laboratory, the discovery of cures, and the care of the sick. Important, too, is her disregard for pursuits which require persuasion or selling, forms of writing, reporting, or literary criticism.

## B. THE PRESENT STUDY

A priori consideration of the importance or lack of importance of the characterstic of dominance in the personality of the nurse resulted in the

---

*Received in the Editorial Office on February 13, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was subsidized in part by a grant from the Research Committee of Santa Barbara College, University of California.

decision to include the *A-S* Reaction Study (1) in the battery of tests given to applicants to a college of nursing. The present study was undertaken to determine the value of this particular scale as a means of identifying the successful nurse or more specifically to determine the success of the scale as a discriminating measure in separating this group from a control group of liberal arts education majors. The experimental and control groups were those used in the previous study which investigated the Kuder Preference patterns. Eighty students in the Knapp College of Nursing in Santa Barbara, California, ranging in age from 17 to 25 years, all Caucasian with the exception of one Japanese girl, made up the experimental group. Matched individually for sex, age, percentile on *ACE,* and race, the control group of 50 girls was selected from liberal arts college girls with an education major enrolled in the University of California, Santa Barbara College. Table 1 presents these data. The *A-S* Reaction Study was administered to the student nurse as part of a battery of qualifying tests given prior to admission to Knapp College of Nursing. The education majors took the test on request as one of a short battery of tests.

TABLE 1

MATCHING VARIABLES, EXPERIMENTAL AND CONTROL GROUPS

| | Experimental Group | Control Group |
|---|---|---|
| Variables | N of 80 | N of 50 |
| Age, Mean | 18.7 | 18.8 |
| Age, *SD* | 1.5 | 1.7 |
| *ACE,* Mean percentile | 34.4 | 35.1 |
| *ACE, SD* | 22.6 | 22.7 |

In the scoring system of the *A-S* Reaction Study (1), plus scores denote ascendance, minus values submission, but in the case of women subjects, the scores in the "average" categories (deciles 5 and 6) have plus values ranging from 0 to 10. The total mean scores for both student nurses and education majors fell within the average range, the education major with a mean score of 7.14, earning a decile 5 rating, and the student nurse with a mean score of 2.13, earning a decile 6 rating. The sigmas of the two groups were nearly identical. The chance that these mean scores point to a true difference between the groups is slight, the *CR* of the percentage of the difference being 1.82. These data are to be found in Table 2.

An analytical study of the subjects' responses to the separate items was undertaken to determine their value as a basis for discrimination between the respective groups. Of 126 possible responses to 34 situations, only nine were found in which the *t* value was 2.00 or greater. Five of the nine were significant at the .01 level of confidence. These items are found in Table 3. Examination of these items reveals some rather interesting trends. In response to Item 14, *Have you crossed the street to avoid meeting some person?*, the response, *never* (a +1 or dominant response) is checked by 59 per cent of the student nurses as contrasted with 32 per cent of the education majors. Sixty-eight per cent of the education majors check the response *occasionally* (given a 0 score by the authors), while it is checked by only 33 per cent of the student nurses. These two responses share honors with a third response found in Item 28 in having the highest *t* value in the list of

TABLE 2

MEAN SCORES, *A-S* REACTION STUDY, EXPERIMENTAL AND CONTROL GROUPS

|  | Mean | Sigma | $SE_m$ |
|---|---|---|---|
| Nurses | 2.13 | 15.15 | 1.69 |
| Educ. majors | 7.14 | 15.45 | 2.18 |

TABLE 3

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN RESPONSES OF 80 STUDENT NURSES AND 50 COLLEGE WOMEN TO 9 SITUATIONS IN THE *A-S* REACTION STUDY

| Situation number | Situation and answers with scores | % Checking Exp. | % Checking Cont. | *t* Value of Diff. | l.c. |
|---|---|---|---|---|---|
| 14 | Have you crossed the street to avoid meeting some person? |  |  |  |  |
|  | Occasionally (0) | 33 | 68 | 4.22 | .001 |
|  | Never (+1) | 59 | 32 | 3.10 | .01 |
| 16a | Have you appeared as lecturer or entertainer before gatherings of over ten people? |  |  |  |  |
|  | Never (—3) | 29 | 12 | 2.60 | .01 |
| 16b | Have you experienced "stagefright"? |  |  |  |  |
|  | Occasionally (—1) | 46 | 68 | 2.60 | .01 |
| 18 | In playing games when young, did you take the lead and decide what the group should play? |  |  |  |  |
|  | Usually (+2) | 4 | 26 | 2.44 | .02 |
|  | Seldom (—3) | 23 | 10 | 2.10 | .04 |
| 23b | Are you embarrassed if you have greeted a stranger whom you have mistaken for an acquaintance? |  |  |  |  |
|  | Very much (—3) | 37 | 18 | 2.53 | .02 |
|  | Somewhat (0) | 61 | 78 | 2.21 | .03 |
| 28 | If you have been in accidents or fires at school or elsewhere, where there were several persons present, have you as a rule taken an active part in assisting? (+1) | 52 | 41 | 3.16 | .01 |

nine. In Item 28 to the question, *If you have been in accidents or fires at school or elsewhere, where there were several persons present, have you as a rule* permits one of three choices: *taken an active part in assisting* (+1), *taken the part of spectator* (—2), *left the scene at once* (0). Eighty-two per cent of the education majors checked the first choice as contrasted with 58 per cent of the student nurses. Since this has a +1 value, our nurses appear to be more submissive in this situation. In answer to Item 16a, *Have you appeared as lecturer or entertainer before gatherings of over ten people?,* 29 per cent of the nurses but only 12 per cent of the education majors checked the answer, *never,* an answer which rates a —3 score, supposedly highly indicative of submission. Somewhat paradoxically, answers to the next item, 16b, *Have you experienced stage fright?,* show a higher percentage of education majors than nurses answering *occasionally.* This may be a function of the lack of experience of the nurse in appearing as an entertainer as indicated in answers to the previous question mentioned above. Only four per cent of the nurses, as contrasted with 26 per cent of the education majors, admit to *usually* taking the lead in games when young. Twenty-three per cent of the nurses reported that they *seldom* took the lead while 10 per cent of the education majors gave the same response. The *t* values for these two items are 2.44 and 2.10 respectively. Thirty-seven per cent of the nurses in responding to the question, *Are you embarrassed if you have greeted a stranger whom you have mistaken for an acquaintance?* checked the response, *very much,* as against 18 per cent of the education majors. However, 61 per cent of the nurses answered *somewhat* to the same question as contrasted with 78 per cent of the control group.

In four of the nine responses the student nurse chose the submissive answer. In two instances the neutral response was chosen by a smaller percentage of student nurses than education majors. In the remaining three situations, the student nurse checked the ascendant response.

The number of items which differentiated these groups at the .05 level of confidence or better is too small to permit conclusions to be drawn, but more significant than limitations in numbers is the lack of direction or trend in the findings. Suggested in the data is the possibility that these groups are not differentiated on the basis of ascendance-submission. More accurate in terms of the findings of this study is the statement that dominance, as measured by the *A-S* Reaction Study, does not discriminate the student from the education major.

This point is further emphasized in the comparison of total mean scores of the two groups and their respective sigmas. The mean scores fall into adjacent deciles, the nurse being slightly less ascendant than the education major, but both deciles are identified as "average" by the authors. The sigmas are nearly identical. Again we find no distinctive pattern which characterizes either group.

## C. SUMMARY AND CONCLUSIONS

1. An investigation was undertaken to determine the value of the *A-S* Reaction Study in identifying the successful nursing college student. The basis for evaluation was the success of the scale in discriminating between a group of student nurses and a group of liberal arts students majoring in educational curricula. A total of 80 Knapp College of Nursing students were matched for sex, race, age, and percentile on *ACE* with 50 education majors from the University of California, Santa Barbara College.

2. Total mean scores for both student nurses and education majors fell within the "average" range. The education major with a mean score of 7.14 and a decile 5 rating might be assumed to be somewhat more ascendant than the student nurse with a mean score of 2.13 and a decile 6 rating. The chances of this discrepancy representing a true difference is slight, the *CR* of the difference in means being 1.82.

3. Further analysis of the responses to the individual test items as a basis for discrimination was not particularly fruitful. The number of responses which were significant at the .05 level of confidence was nine; five of these were significant at the .01 level of confidence. The trend was toward a submissive pattern on the part of the student nurses in four responses. The education majors outnumbered the student nurses in the choice of two items, each with a zero rating, indicating a neutral position. In the remaining three items, the student nurse assumed the more dominant position.

4. The number of significant items in this inventory is too small to make any valid claim as to greater or less dominance value as measured by the *A-S* Reaction Study in the student nurse as compared with her sister in the liberal arts education curricula. Neither comparison on the basis of total scores nor on the basis of individual choices yields highly discriminatory data. We can only conclude that the two groups do not show a true difference in respect to the quality of dominance as measured by the *A-S* Reaction Study.

## REFERENCES

1. ALLPORT, G. W., & ALLPORT, F. H. The A-S Reaction Study: A Scale for Measuring Ascendance Submission in Personality. Manual of Directions, Scoring Values, and Norms. Boston: Houghton Mifflin, 1939. (Pp. 2-16.)

2. Beaver, A. P. Personality factors in choice of nursing. *J. Appl. Psychol.*, 1953, **37**, 374-379.

3. ———. Kuder interest patterns of student nurses. *J. Appl. Psychol.*, 1953, **37**, 370-373.

*Department of Psychology*
*University of California*
*Santa Barbara College*
*Santa Barbara, California*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 79-82.

# MISCONCEPTIONS CONCERNING THE INTERPRETATION OF SUB-GROUP VARIATIONS WITHIN NORMATIVE DATA*

*Utah State Hospital*

LEWIS AUMACK

In a recent paper by Applezweig (1) *MMPI* responses of a sample of Denver high school and college students "differing from the standardization population in regard to education and assumed to be different in regard to intelligence" were collected and compared with the original standardization data. Based upon the finding that significant differences existed for 8 of the 9 scales for the males and 5 of the 9 scales for the females, it was concluded that "the 'general normal group data' do not apply for all scales" and therefore argued that (*a*) "more adequate norms" are required and (*b*) "utmost caution must be exercised in the interpretation of profiles of persons in advanced school situations." The present set of notes are herewith set forth in an attempt to clarify fairly common, yet serious, misconceptions concerning interpretations of sub-group variations within a normative framework, and then to offer what appear to the writer to be more valid conclusions concerning such variations.

Taking an example of such a normative test as the *MMPI,* the evaluation of any particular individual's profile may proceed, on the one hand, in relation to the more general standardization norm or, on the other, in relation to the norms of select psychological or sociological subgroupings within that general norm. If we assume, for example, that a clinician gets his hands on a given student whose profile approximates the mean profile of Applezweig's sample of lowest *IQ* males, he would note, *in comparison with the standardization norms,* such deviations as the generally increased elevation of the total profile; the relatively elevated F, Pd, Mf, Sc, and Ma; the relatively over-controlled Pa; the tendency toward Sc>Pt; Anxiety Index of 53 and Internalization Ratio of .94 (7); and so forth. On the basis of such observations, the clinician might interpret such a person has difficulty in relating to others on other than a superficial basis due to his inability to perceive realistically his own rôle-playing deficiency; experiences feelings of social alienation; exhibits an inability to withstand frustrations of marked

---

*Received in the Editorial Office on February 15, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

narcissistic needs for but brief intervals of time; and due to his underlying hypersensitivity concerning immediate authority situations, dispels the tensions of such frustrated needs by acting-out more-or-less overtly against more distant and generalized authority figures. Furthermore, such a clinician might note that the patterns for the lowest age, educational, and intellectual groupings in Applezweig's study are similar enough to warrant the assumption of a more general dimension of *maturity-immaturity* (an assumption partially borne out by the intercorrelation of +.72 between age and educational level and the probability of a significant positive correlation between these and intellectual functioning). In discussing the type of profile represented by the lowest *IQ* grouping, therefore, he would be discussing but one aspect or manifestation of an immature personality—which, in a number of ways, is the essential feature of the preceding interpretation. It would appear, then, that a sub-group variation may be used to validate rather than invalidate the use of the standardization group as a reference point for certain types of interpretation.

Another equally valid approach for certain other purposes would be to compare a given individual's profile with the norm of whatever sub-group or groups to which he belongs or is identified. In this instance, the clinician might complicate the interpretation with such comments as, "This student's Sc T-score of 60 (in the context of other scale scores) indicates that he is more sensitive and withdrawn than the average person but that in comparison with others his own age (or intellectual and educational level) he is relatively insensitive and out-going." Although the discussion so far has been devoted to an examination of the *MMPI,* it seems reasonable to assume that the same comments could be made concerning the profiles of such other normative tests as the Wechsler-Bellevue, Rorschach and others. Certainly no one finding that a particular occupational group (e.g., auto mechanics) had either total *IQ* scores or profile patterns that deviated significantly from the standardization norm would therefore conclude that "the general normal group" data did not apply, or that one should be "cautious in interpreting" such deviations. Rather, one would say that a given auto mechanic's total score, sub-test score, or pattern of scores differs from the general populace in such and such a manner or degree; furthermore, that such test responses do or do not differ in such-and-such a manner or degree from other sub-group norms (clinicians, physicians, bachelors, or bird-watchers). In other words, the uncovering of sub-group variations within normative data may function to refine the test for further and more complicated relativistic evaluations.

The comments made so far seem to me to be rather obvious and but lay the groundwork for the tackling of a more subtle and significant problem, that of conceptualizing differential criteria of effective functioning. Essentially the question raised by such data as Applezweig's is this: "Here we have high school and college students going to school and staying out of the 'booby-hatch' and yet they attain scores on the *MMPI* that deviate significantly in the direction of pathology. How come?" Similarly, the occasional clashes between overly-staunch Rorschach-ers who state that the *MMPI*'s of this or that group of normals represented psychotic profiles, and conversely *MMPI*-ers who claim that pathology can be found in any Rorschach, appear to represent a more extreme expression of the same query. My own feeling is that the answer is not to be found as much through increasing differentiations into various and sundry sub-group norms as it is through the dispelling of the naïve propensity for equating norm deviations with pathology with the opposing point of view that certain trait deviations may work *for* rather than *against* effective functioning for certain psychosocial groupings. On the basis of the presently discussed study, e.g., it would appear that elevations on *D* for male, and lowered Pt for female, graduate students reflect types of personality structuring that leads to above-average functioning in an intellectual milleau, while for a different group, e.g., psychotherapists, similar deviations might possibly be indicative of personal and/or social malfunctioning.[1]

My own overall bias, therefore, is that insofar as misinterpretations of "pathology" are likely to result from practically any psychological test, the most valid solution would seem to lie not only in the differentiation of general norms into smaller and more homogeneous standardization groupings, but also in the re-examination of our traditional absolutist and intrapsychic formulations of pathology which, until recently, have neglected differential and relativistic social criteria of adjustment and effective functioning (2, 4, 6). Until such differential criteria of "normality" are conceptualized, explicitly formulated, and configurally evaluated, my own "utmost caution" in interpreting such infernal test devices is going to be exceeded only by my caution in taking them!

---

[1] This is even more likely to be the case when it is noted that scores obtained from empirically derived scales, particularly such complex ones as represented on the *MMPI*, are not necessarily homogeneous measurements but may be the summations of psychologically distinct sub-scale groupings of items (defined either rationally (3) or empirically (5)) having differential loadings for normal and clinical samples. If on certain sub-scales, e.g., the elevated Pa scores of Applezweig's female subjects were most heavily saturated with items dealing with *moral virtue and naivete*, whereas equally high Pa scores of institutionalized patients resulted from saturations of items dealing with *ideas of external influence*, one would be less likely to assume pathology with the former than with the latter.

## REFERENCES

1. APPLEZWEIG, M. H. Educational levels and Minnesota Multiphasic profiles. *J. Clin. Psychol.,* 1953, **4**, 340-344.
2. GOUGH, H. G. Predicting success in graduate training: A progress report. Berkeley: Univ. California, Institute of Personality Assessment and Research, 1950. (Mimeographed.)
3. HARRIS, R. E. Subscales for the *MMPI.* San Francisco: Univ. California, Langley Porter Clinic, 1949. (Mimeographed.)
4. KELLY, E. L., & FISKE, D. W. The Prediction of Performance in Clinical Psychology. Ann Arbor: Univ. Michigan Press, 1951.
5. LITTLE, K., & FISHER, J. The internal structure of the *MMPI* Hysteria Scale. Paper read at Bay Area Veterans Administration Clinical Psychology Research Meeting, October, 1952.
6. OSS, ASSESSMENT STAFF. Assessment of Men. New York: Rinehart, 1948.
7. WELSH, G. S. An anxiety index and an internalization ratio for the *MMPI. J. Consult. Psychol.,* 1952, **16**, 65-72.

*Utah State Hospital*
*Provo, Utah*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 83-87.

# THE EFFECT OF "PUNISHMENT" (ELECTRIC SHOCK) ON FIGURE-GROUND PERCEPTION*

*Departments of Psychology, University of Michigan and Cornell University*

---

DONALD E. P. SMITH AND JULIAN E. HOCHBERG

---

## A. THE PROBLEM

The effect of reward and punishment on perception as reported by Schafer and Murphy (6) has been contested both on empirical and on theoretical grounds (4, 5, 7).

Their study used two sets of two half-moon "faces," the two faces of each set forming a circular reversible figure-ground when joined at the profile (6, p. 336). Subjects won four pennies when each of two of the single faces was projected tachistoscopically, and lost two pennies when each of their companion faces were presented. In a test series, the two faces of each set were combined so that either could be seen as figure. Rewarded faces were reported more often than punished ones, suggesting the operation of autism (and of pre-perceptual recognition, 1) since autism as obtained here logically seems to require that, *at some level prior to figure-ground formation,* both faces are "perceived" and reacted to in terms of the rewards and punishment of the training series.

Wallach (7), assuming that an area must be organized into figure and ground before trace arousal and recognition can occur, suggests a possible explanation: the *outline* figures used by Schafer and Murphy permitted recognition of "line" rather than of figural form, and the form of such a profile "line" may remain recognizable in both figures, unlike the *contour* of a solid figure; he suggested, therefore, that repetition with *solid* figures, would not yield "autism." Rock and Fleck (5), attempting replication of the Schafer and Murphy procedure, using *outline* figures, found no consistent results, but (3) the rewards used (two and four cents) may not have been motivationally equivalent in 1943 and 1950.

The first problem is whether, with electric shock as punishment rather than questionable monetary rewards, "perceptual defense" can be obtained with *solid* figures. The problem is similar to that of McLeary and Lazarus (2) although "faces," rather than nonsense syllables, were used.

---

*Received in the Editorial Office on February 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

## B. EXPERIMENT I

### 1. *Method*

Two solid white (*B, D*) and two solid black (*A, C*) "faces" identical in contour to those of Schafer and Murphy (6, p. 336) but with no "outlines," were reproduced on 2" x 2" slides, singly, for the training series and in two combinations (*A-B, C-D*) on a gray ground for the test series, and projected to 15 inches in diameter. Two groups of 10 subjects each were seated at a table, five at a time, separated by cardboard blinds; their distance from the screen varied from 84 inches to 96 inches. Electrodes were attached to their hands and to induction coils and 6-volt dry cells in single series to a telegraph key operated by *E*.

Fifteen 1/3 second exposures were made of each of the four faces in random order. Electrical shock accompanied one member of each pair each time it was presented. Before each presentation during the first 12 trials, attention was directed to the appropriate one of four cardboard "help" faces mounted on the wall (non-reversible, "improved" figures), and the name (*A, B, C,* or *D*) was given. To aid learning, *S*'s called out the name of each figure after presentation for the remainder of the series.

Group I was shocked simultaneously with the presentation of Figure *A* and with presentation of *D*, Group II, with *B* and *C*, in order to control possible structural, positional, or color dominance. Thus, if both groups subsequently displayed a significantly greater perception of unshocked than shocked faces, or one group showed this effect significantly while the other group did not differ from it significantly in the *opposite* direction, such results could not be attributed to structural, positional, etc., factors, and the effects would be assumed as due to the shock procedure.

A learning check test of eight presentations, two of each single face, followed the training series.

As in previous studies, a post-training series consisted of 63 presentations (1/3 sec. each): 16 of each of the two reversible drawings (*A-B, C-D*) and 31 of two extraneous "set breaking" figures, irreversible faces pointing right or left. After each test figure, the extraneous figure pointing in the direction opposite to that of the non-shocked face was presented.[1] Subjects wrote the name of the face seen, "X" for extraneous figures and "?" if the face could not be identified.

---

[1] This procedure differs from the previous studies in which the extraneous figure was determined by the preceding response of the subject (impossible with groups of subjects); however, the results showed no directional effects following the set-breaking figures ($P > .10$).

## 2. *Results*

For Group I, the difference between responses of shocked and non-shocked faces is not significant (Table 1); that for Group II *is* ($P < .01$), as are the pooled results ($P < .01$).[2]

Analyzing each set, *A-B* and *C-D*, alone, Rock and Fleck reported a significant tendency for *S*'s to report punished faces in the *A-B* situation and to report rewarded faces in the *C-D* situation. In the present case, 164 unshocked and 101 shocked faces were reported for the *A-B* figure, and 98 unshocked and 81 shocked for the *C-D* figure, i.e., the same direction in both cases. In short, while purely "figural" factors were of importance, the shock procedure clearly had effect.[3]

TABLE 1

NUMBER OF CORRECT RESPONSES TO POST-TRAINING TEST FIGURES

| Group | No. of subjects | NS | S | P | B |
|---|---|---|---|---|---|
| | | *Experiment I* | | | |
| I | 10 | 116 | 128 | | 1 |
| II | 10 | 146 | 54 | $<.01$ | 0 |
| Sum: | | 262 | 182 | $<.01$ | 1 |
| | | *Experiment II* | | | |
| III | 5 | 41 | 33 | | 75 |
| IV | 5 | 28 | 26 | | 32 |
| Sum: | | 69 | 59 | | 107 |

*Note: NS, S* list the number of correct responses to the non-shocked and shocked figures, respectively; *B* is the number of correct identifications of both faces; *P* is the probability of obtaining such proportions by chance.

However, although something which may be called "perceptual defense," was obtained, this still does not necessarily imply pre-perceptual (or even pre-figural) recognition. Despite Wallach's belief that with solid figures *S*s would not be able simultaneously to recognize both "faces," this may be just what occurred. Hence, the next problem is to determine whether, with appropriate instructions, *S*s are capable of identifying both faces (if this is the case, of course, we would also expect an attendant decrease in or absence of "autism" in these circumstances).

---

[2] Rock and Fleck discarded the results of subjects who failed to recognize all faces in a check test administered after the experiment proper; we here retained all subjects. No significant differences ($P > .10$) in "autism" were found between those subjects who made one or more errors on the check list and those with perfect scores.

[3] It seems unlikely, moreover, that the "autism" (at least, that obtained in this experiment) is due to inhibition of the verbal "naming" response, since the errors in identification did not contain significantly more shocked than unshocked responses ($P > .10$).

## C. EXPERIMENT II

### 1. *Method*

Ten *S*'s were trained as in Experiment I, five in Group III and five in Group IV. To determine whether it is possible to identify *both* figures in a given 1/3 sec. test presentation, *S*'s received written directions prior to the post-training series to the effect that two faces would appear in juxtaposition at each presentation and that they should attempt to identify both.

### 2. *Results*

Nearly half (107) of all correct reports (235) identified both stimulus figures (Table 1). Here the difference between shocked and unshocked faces is not significant, but the increase in correct reports of *both* faces, as compared to Experiment I, is ($P < .01$).

## D. DISCUSSION

The results of the first experiment do not contradict the original findings of Schafer and Murphy. Wallach's expectation that "autism" would not appear with solid figures was not fulfilled. Apparently the "punishment" (shock) procedure did affect the selection of the figure perceived.

But does the individual first perceive (albeit unconsciously) the painful stimulus so that he may avoid perceiving it? Actually, we need not assume that figure-ground organization springs into all-or-none dichotomy, even in 1/3 second. If, during the *early* stages of differentiation, before perceptual integration is complete (cf. Murphy and Hochberg, 3), *one portion of a given boundary may serve as contour for the black region while in another segment, the white is "figure,"* two not-necessarily exclusive explanations suggest themselves:

1. Wallach's objection to the use of outline figures may be a more general one, and may apply to solid figures as well, since it appears from the second experiment that it is possible to identify aspects of both figures, and, thereby, a "choice" between them based upon Effect is at least conceivable.

2. Shock may be a "distracting" deterrent to perceptual learning (or "integration"). The unshocked figure, perhaps with briefer latency, might be favored in the integration of the final figure-ground organization; several *S*'s reported that the shock "made it harder to learn the faces" or "took my attention away from the faces."

## E. SUMMARY

The contested "autism" effect of Schafer and Murphy was tested with solid (rather than outline) figures and with electrical shock as "punishment." With 20 *S*'s a significant difference favoring non-shocked figures was obtained, against Wallach's prediction; however, his objection to outline figures may actually hold here as well since a second experiment with 10 *S*'s indicated that with the proper set the identification of *both* phases of a reversible figure-ground configuration can be increased, suggesting explanations which do not invoke pre-recognition processes such as "subception" (2).

## REFERENCES

1. HOCHBERG, J., & GLEITMAN, H. Toward a reformulation of the perception-motivation dichotomy. *J. Personal.*, 1949, **18**, 180-191.
2. MCCLEARY, R. A., & LAZARUS, R. S. Autonomic discrimination without awareness: an interim report. *J. Personal.*, 1949, **18**, 171-179.
3. MURPHY, G., & HOCHBERG, J. Perceptual development: some tentative hypotheses. *Psychol. Rev.*, 1951, **58**, 332-349.
4. PASTORE, N. Need as a determinant of perception. *J. of Psychol.*, 1949, **28**, 457-475.
5. ROCK, I., & FLECK, F. S. A re-examination of the effect of monetary reward and punishment on figure-ground perception. *J. Exp. Psychol.*, 1950, **40**, 766-776.
6. SCHAFER, R., & MURPHY, G. The rôle of autism in a visual figure-ground relationship. *J. Exp. Psychol.*, 1943, **32**, 335-343.
7. WALLACH, H. Some considerations concerning the relationship between perception and cognition. *J. Personal.*, 1949, **18**, 6-13.

*Department of Psychology*
*University of Michigan*
*Ann Arbor, Michigan*

*Department of Psychology*
*Cornell University*
*Ithaca, New York*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 89-95.

# SOME ANTHROPOMORPHIC COMMENTS ON LATENT LEARNING*

*Department of Psychology, University of Washington*

ALLAN KATCHER

## A. THE FUNCTION OF ANTHROPOMORPHISM

The term anthropomorphism is usually applied to any attempt to ascribe human-like attributes to animals, or in a broader sense, to interpretations of the behavior of other organisms in terms of the observer's own. Psychologists are critical of any signs of anthropomorphism which are detected in scientific discussions, as indeed they should be, if the "guilty" person is convinced that he has explained animal behavior merely by his analogous equation of animal and human processes. However, thinking anthropomorphically, we can and often do develop hypotheses which lead to the discovery of variables previously neglected or regarded as insignificant. Still another advantage of such thought lies in the resultant sensitization to possible human analogues of problems encountered in animal psychology. It is, after all, the former level that is the focus of most of our interest, if not our research.

## B. THE COMMENTS

Many discussions of learning begin by stating that it is not observed directly but is inferred from performance. Thus behavior changes in response to the "same" stimulation, assuming they are not due to fatigue, maturation, etc., are taken as evidence that learning has occurred. In numerous instances the changes are gradual (e.g., errors or time decreases, the number of responses increases) and, with respect to a given criterion, a learning course or trend can be described. Latent learning is certainly subsumed in these discussions but appears to be "latent" or *undiscovered* when the gradual trend usually observed is not evident in performance. In a recent review of the literature, Thistlethwaite distinguished four types of latent learning studies:

Type I, or the Blodgett variety, which consists in giving rats a series of unrewarded, or slightly rewarded, trials in a maze. A relevant goal object is then introduced and additional trials are given. Sudden or marked drops in the

---

*Received in the Editorial Office on February 17, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

number of errors exhibited by the previously unrewarded group, as contrasted with a group of consistently rewarded animals, has been considered evidence of "latent learning."

The second type is characterized by the procedure of permitting the animals to explore and live in the empty maze for brief periods of time. Subsequently a relevant goal object is introduced into the maze, and a single test trial or series of test trials are given.

The third consists in giving rats which are satiated for food and water a series of trials in a maze the pathways of which contain the goal objects for which the animals are satiated. After this series of trials the rats are made hungry or thirsty and given additional free trials in the maze.

The fourth involves giving thirsty or hungry rats a series of trials in a maze containing relevant and/or irrelevant goal objects. The animals are then satiated for the formerly desired goal object and deprived of the previously undesired goal object. One or more trials are then given in the practiced maze. Many investigators have failed to find evidence of latent learning with this type whereas positive results have been reported by others (9, p. 99).

Experiments of Types I and II pose a serious problem to anyone committed to the view that performance must reflect continuous and gradual changes before learning criteria are reached, i.e., that each successive experience provides an additive unit to acquisition. Latent learning results are especially challenging if one further assumes that performance changes cannot occur without "relevant" reinforcement. The "saving explanation" offered by the "gradual and reinforcement" learning theorist states that the latent learning experimenter has not considered the possibility that the animal may run the maze under the influence of a drive (or drives) different from the one assumed to be relevant, which can be reinforced by the experimental setup. Thus reinforcement of an exploratory drive may occur during the "unrewarded" trials. Another hypothesis of similar vintage is more anthropomorphic, in the detrimental sense. This one maintains that the reason why maze-wise rats, as well as naïve rats, sometimes manifest learning prior to the introduction of food is that the rats were really rewarded by removal from the maze, even though, in the case of the experimental group, food-reward was not presented immediately upon completion of the maze. Removal of the maze is thus regarded as rewarding in itself. Do these theorists mean that rats perceive maze running as undesirable? If so, then we should find it interesting to contrast the performance of maze-reared rats and cage-reared rats when tested in this type of situation.

In some experiments (Type III) the animal is required to associate the locus of a non-relevant reward at the first time the "drive" under which he is presumed to be running is shifted. From a reinforcement viewpoint, the previously reinforced behavior should interfere with the shift, and the prediction made that the animal will continue to run to the side that was reinforced until extinction occurs. This of course, assumes that it is the response to the appropriate goal which is learned, and not, as expectancy theorists might predict, the route. If the latter were the case then three things would be required before latent learning were exhibited: (*a*) drive discrimination, (*b*) path discrimination, (*c*) the association of the appropriate route with the ongoing drive. The available evidence in the literature is by no means clear as to whether latent learning does occur in this instance.

Let us turn for a moment to a more familiar subject. On the college level students frequently provide eloquent testimony on these "learning" problems whenever they complain that examinations fail to measure what they "got out of the course." The student may mean that *the test did not measure what he had learned,* that is, failed to sample the areas in which he was prepared, or the test was so constructed that detailed knowledge of the subject led him to select an alternative which *the instructor did not regard as correct.* Conversations with students are quite revealing to the instructor who believes his tests provide valid measures of learning. Their experience would suggest that learning, as students define it, may result in the same level of performance because of the insensitiveness or unfair tests which the instructor has employed. An experiment by Postman and Tuma (5) provides a real illustration of this point. Using a verbal maze with alternative pathways they found that one experimental group of subjects, trained on only one pathway recalled almost half of the stimuli marking an alternative but irrelevant pathway, while another experimental group, tested for transfer of incidental information to the learning of the alternative route, failed to evidence any facilitative benefits. The authors criticize the tendency for many investigators to state the latent learning hypothesis merely in negative terms and argue for research based upon statements dealing with the cognitive changes which are expected to result from the original training. In addition they hypothesize that ". . . if rewarded training is for an organized sequence of responses to a series of stimuli, there will be incidental learning of the presence, but not of the serial arrangement, of the irrelevant stimuli" (p. 123).

## C. The Neglect of Individual Differences

In experiments with human subjects attention is usually focused on variables which have been largely ignored in most animal studies of latent learning. Both Expectancy and Reinforcement theorists have conducted their studies as if individual differences in animals were wholly irrelevant to the problem under consideration. Yet in the analogous human problem we would have attempted to control the so-called learning ability or "intelligence," age, and prior experience of our human subjects. Otherwise it would be impossible to determine whether any differences between the two groups could be attributed to experimental conditions or to these other subject differences. One facetious comment that has been made frequently holds that California rats seem to behave one way and Iowa rats another. Perhaps it is time for someone to compare systematically the performance of the two strains in a wide variety of learning situations.

In raising this point I am impressed by some other human observations. One thing characteristic of extremely intelligent people is their possession of large stores of detailed information about relatively irrelevant material which they constantly integrate and use in a wide range of contexts. Facility in manipulating information and detailed knowledge seemingly go hand in hand although they are not perfectly correlated. To take another example, we are all familiar with the person who knows how to get to places, but who cannot explicitly direct anyone else to the place. On a verbal test of spatial localization the hypothetical man (a cab driver, perhaps) might easily seem to be lost, but on a performance test his success would be noteworthy.

It is possible that reinforcement theories apply to certain situations and to certain animals—that for other situations, different principles have to be employed, although they could conceivably be derived from reinforcement theory as in the case of Seward's response surrogates (7). At the human level we are acquainted with the kind of student who can only acquire knowledge by constant repetition. Fortunately we also meet students who readily grasp knowledge, retain it well, and utilize it in contexts far afield from the original learning. At a certain level of performance criteria can be reached by both processes, but there would be a considerable difference in the length of time required to learn. Similarly, there are numerous instances of child behavior where "new" behavior is exhibited which depends upon previous experience but has never been deliberately reinforced.

How about corresponding behavior in the rat? Could not some of the discrepancies in latent learning studies be accounted for by individual differences in subjects? For example, if dull rats were employed in the typical latent learning situation they might exhibit the shift in response only when continually and consistently reinforced for a given response; the brighter rats would evidence latent learning. Possibly it is not intelligence alone, but intelligence plus other factors such as exploratory drive, overall activity, etc., which would have to be taken into account. It is difficult to answer the above question, for, according to Christie (1) there is considerable variability in the strains of rats used and few relevant comparative studies to permit an adequate evaluation of this hypothesis. In addition, animal investigators have been all too cavalier in reporting details about their experimental animals, often ignoring such factors as strain, age, sex, previous maze experience, manner of handling, etc. Many studies report dropouts of subjects whenever they fail to learn in the initial experimental stages, with the proportion rising as high as one-third or more in some experiments. If there are individual differences in rats then such a procedure may mean that findings about latent learning, as well as other kinds, hold only for very restricted populations. One attempt to circumvent this problem would be to assign animals randomly to experimental and control groups as in most human studies. However, one could conceivably argue, apropos of our previous human discussion, that if one began with a "dull" strain it would still be impossible to obtain evidence of latent learning.

Certainly there are relevant experimental data on individual differences in rats' maze performances. Tryon (10) was successful in breeding two relatively distinct groups of maze learners—"brights" and "dulls." Searle (6) found that these groups differed primarily in multiple maze situations where food served as the incentive. In another study, Krechevsky (3) obtained evidence that the "brights" exhibited significantly more spatial hypotheses than the "dulls." The latter finding could be quite significant for latent learning experiments. In studies of reasoning in rats, Shepard (8) observed that some rats exhibited consistently superior performances in reasoning, although these were not correlated with performance on unit-alike and unique-unit mazes. Other investigators (2, 4) have similarly reported instances of animals who were particularly adept in some performances and not in others. Presumably other variables such as activity differences, exploratory behavior, innate preferences for certain kinds of stimuli, etc., also vary and interact with "intelligence" in rats.

There has been a recent surge of interest in pre-experimental differences in animals which I hope will be paralleled by studies of ability differences. Christie's article (1) emphasizes that many factors in the rat's pre-experimental experience can influence his maze performance. Such variables as familiarity, handling, accidental deprivation, and maintenance schedule all appear to be relevant. Too often we have blithely assumed that the rat is an empty organism, or that all rats are filled with the same sort of kapok, but these trends indicate that we have some extensive ground-work to do before we can have confidence in the generality of our present theoretical statements, as well as the fact that the rat is not as ideal an experimental animal as we once considered him.

## D. Back to Humans

Since attempts to construct intelligence tests for animals entail numerous difficulties it might be more fruitful to shift our research to experimental analogues of latent learning with human subjects. Postman and Tuma's (5) verbal maze could with some modifications provide a framework for the evaluation of hypotheses about individual differences, especially with regard to such factors as set, interest, comprehension of task, sensitivity to stimuli, as well as intelligence. Incidentally, their concern about the nature of the learning that takes place in latent learning suggests a parallel type of inquiry into the nature of factors affecting intelligent behavior, a topic which has not been appreciably advanced by the conventional substantive approach to mental testing.

Another analogue was suggested by an observation of children playing Authors, a game consisting of playing cards on which names and brief biographical sketches of authors are printed. It is based upon the acquisition of cards whose names are similar to those held in the other players' hands. At the conclusion of the game the girls were asked to identify the authors of several books and plays. The "brighter" ones correctly identified them, despite the fact that the game itself paid no premium for learning this information, and over the long course of time differential experiences would have equalized the associative success and failure values. Unfortunately, no intelligence data was available to bear light on the nature of the individual differences and the above comments on individual differences simply represent clinical hunches. Similar experiences could provide interesting experimental paradigms for future work in child psychology, in which some of these aforementioned variables could be systematically investigated.

It is obvious that we all acquire "irrelevant" information—some through principles of perceptual organization, others through exploratory orientation and experience. However, whether we utilize such information is dependent not only upon association, significance, expectancies, and reinforcement, but also upon native factors as well.

By way of conclusion I ought to remark that I have ranged a little far from latent learning proper only because the problems of latent learning reflect the problems of psychology at large: the significance of the variables we observe, measure, and relate.

## REFERENCES

1. CHRISTIE, R. Experimental naïveté and experiential naïveté. *Psychol. Bull.,* 1951, **48**, 327-339.
2. HERON, W. T. Inheritance of maze-learning ability in rats. *J. Comp. Psychol.,* 1935, **19**, 77-89.
3. KRECHEVSKY, I. Hereditary nature of "hypotheses." *J. Comp. Psychol.,* 1933, **16**, 99-116.
4. LOEVINGER, J. "Reasoning" in maze-bright and maze-dull rats. *J. Comp. Psychol.,* 1938, **26**, 413-426.
5. POSTMAN, L., & TUMA, A. H. Latent learning in human subjects. *Amer. J. Psychol.,* 1954, **67**, 119-123.
6. SEARLE, L. V. The organization of hereditary maze-brightness and maze-dullness. *Genet. Psychol. Monog.,* 1949, **39**, 279-325.
7. SEWARD, J. P. Experimental evidence for the motivating function of reward. *Psychol. Bull.,* 1951, **48**, 130-149.
8. SHEPARD, J. F. Higher processes in the behavior of rats. *Proc. Nat. Acad. Sci.,* 1933, **19**, 149-152.
9. THISTLETHWAITE, D. A critical review of latent learning and related experiments. *Psychol. Bull.,* 1951, **48**, 97-129.
10. TRYON, R. C. Genetic differences in maze-learning ability in rats. In: *Nat. Soc. Stud. Ed., YrBk,* Pt. I, 1940, 11-119.

*Department of Psychology*
*University of Washington*
*Seattle 5, Washington*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 97-101.

# ONE- AND TWO-SIDED STATISTICAL TESTS AND OBJECTIVES OF EXPERIMENTATION*

*Department of Psychology, University of Denver*

A. B. Shaklee

The alternatives of using one- or two-sided tests of significance have been recognized for some time. In certain cases, the nature of the test being used has automatically limited the choice to one alternative. This is true of one-sided $F$ tests in the analysis of variance. Paradoxically though, the logic has been more elusive with some of the simpler tests of significance involving one or two samples. Several modern textbooks in statistics show inconsistencies from one to the next concerning reasoning directed toward this problem.

Recent articles on the subject (1, 3, 4, 6, 7) have called attention to certain distinctions that are important for the experimentalist. Unfortunately, however, in a number of these articles suggestions are made which overlook or misinterpret existing analyses, which stem in large part from the work of Neyman and Pearson. There exists a need for relating current discussion more specifically to careful interpretations available in the literature (e.g., 2, 5, 8). There are also a few points at which it seems desirable to amplify on available treatments in terms of the relation between the statistical rationale and the nature of the problems to which the tests are applied. In one of the recent articles, Marks (6) was interested in classifying types of experimentation according to the nature of the statistical test used. He states: *"Experimentation$_I$ is that form of experiment to which two-tailed tests of the data are appropriate; experimentation$_{II}$ is that form of experiment to which one-tailed tests of the data are appropriate."*

"It follows that (*a*) preëxperimental predictions and (*b*) use of one-tailed tests are severally necessary and together sufficient for experimentation$_{II}$" (6, p. 182). Hick (3) correctly points out the inadequacy of a preëxperimental prediction as a justification for a one-tailed test. In his reply Marks says, "All statements made therein [i.e., in his initial article] relative to the testing of hypotheses have the acceptance of statisticians generally" (7, p. 207). This assertion merits further questioning.

Hick goes on to state that if the experimenter knows a negative bias is impossible ". . . he is entitled to put the whole critical region in the positive

---

*Received in the Editorial Office on February 17, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

tail and call it a one-tailed test" (3, p. 317). He later implies that in other than such cases only two-tailed tests should be used. One would like to see such a sweeping recommendation supported by a more detailed analysis of available discussions of the problem.

Burke (1) also presents several sound objections to indiscriminate use of one-tailed tests. Some of his analyses of why preëxperimental predictions concerning the sign of a mean difference are inadequate to justify a one-tailed test seem especially well-done. However, his justification of such tests in one particular situation is likely to create misunderstanding. He states that in many experiments two statistical decisions are to be made. "The first is the decision made by the individual experimenter who frequently plans one experiment from his evaluation of a previous one. We concede that here a one-tailed test is often proper" (1, p. 385). It is true that in a few such cases a one-tailed test would be proper, but it would not be because the experimenter had to use the results of one experiment to plan another. The large majority of situations involving the planning of future psychological research are clearly out-of-keeping with the rationale underlying one-sided tests. There is a real danger that the statement might encourage experimenters to develop hypotheses for new research after ignoring pertinent data or on the basis of a subjective inflation of the meaning of the statistical probability. This danger will be made more apparent by a study of the analysis given below. Burke's later statement that ". . . the adoption of one-tailed tests is equivalent to a general lowering of levels of confidence" (1, p. 386) is true only in terms of certain incorrect usages such as will be described further on, and as a generalization is also misleading.

One criterion involved in selecting the type of test to be used is the power of the tests in question. Power is a function of the Type II error of the test, which in turn is the probability of failing to reject the null hypothesis when it is actually false. In Figure 1 there are schematized, for a particular level of significance and sample size, the power functions of three types of test of an exact hypothesis: the two-sided $t$ test ($P_2$), and the corresponding right-hand one-sided ($P_{1R}$), and left-hand one-sided test ($P_{1L}$). The parameter value for the population sampled is $\mu$, while $\mu_0$ is a specific value against which we test. Not only the power of a test, but also its superiority over another, depends on which alternative is true. Thus $P_{1R}$ is superior to $P_2$ only for situations where $\mu$ is actually greater than $\mu_0$. $P_{1R}$ is very poor for $\mu < \mu_0$. The fact that we wish to test an hypothesis concerning the mean implies that $\mu$ is unknown. It is important for the argument to note that while we may speculate that $\mu_0 - \mu$ will have a particular sign, the true direction of this difference may be the same or the opposite of the predicted direction of difference. We must distinguish here between two kinds of hypotheses—the pre-experimental hypothesis concerning the sign of $\mu_0 - \mu$, and the statistical hypothesis which the overall experimental rationale indicates it is appropriate to test. These two hypotheses may be quite unrelated. The relative power of a particular test varies with the distance and sign of $\mu_0 - \mu$, and not with the hypothesis concerning the sign of this difference. In order to justify the right-hand one-sided test we must not only be operating on the hypothesis $\mu > \mu_0$, but we must also be in the position where it is not important to know if the true situation should be $\mu < \mu_0$, since $P_{1R}$ is unsatisfactory for this alternative.

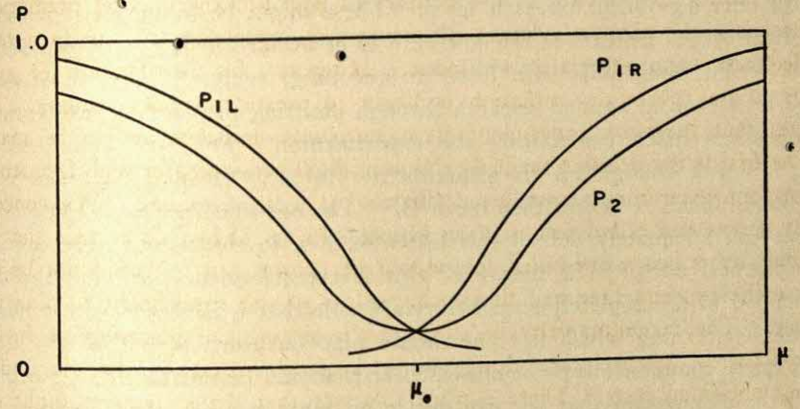


FIGURE 1
COMPARATIVE POWER OF ONE- AND TWO-SIDED TESTS OF HYPOTHESIS: $\mu = \mu_0$

When is it unimportant to us if we commit Type II errors for $\mu < \mu_0$ in the above situation? A commonly cited example has to do with the operating life of light bulbs. A manufacturer may have established with considerable precision that a particular type of bulb now in production has a mean life of a certain number of hours, say $\mu_0$. He wishes to determine on the basis of a sample if a new variety of bulb has a longer mean life. The right-hand one-sided test of an exact hypothesis, $\mu > \mu_0$ can be used here, if it should be of no concern in the event that $\mu < \mu_0$. In certain industrial problems, this latter consideration is often assessed of no consequence. In parallel situations where, say the cost of a new manufacturing process were being compared to the cost of the old, the left-hand one-sided test could be used. In addition to tests of exact hypotheses like the above, it is possible to have comparisons based on two samples where the one-sided test is appropriate. An example may be found in Dixon and Massey (2).[1]

It should be noted that although situations like the light bulb experiment are traditionally cited as being best analyzed by a one-sided test, the two-sided test would be preferable if being able to detect the existence of $\mu < \mu_0$ would offer significant research leads. These might be in terms of possible disadvantageous features of the new bulb to be avoided in future pilot models. Some recent discussions seem likely to encourage inappropriate use of one-sided tests even in applied research through omitting to point out the variety of questions often confronting the experimenter. To justify a one-sided test, it is not enough for the experimenter to say, "It is of no concern to me should mean A be less than mean B." The one-sided test is justified only if one can adequately defend this last assertion, in the light of the overall experimental aims. The nature of this last requirement precludes the development of a completely straightforward rule that will always give a quick answer concerning which test one should plan on using.

A fairly common misuse of one-sided tests should be mentioned. It begins with the assumption that one can decide on what families of alternatives to test *after* one has seen the trend of the results. Then the argument takes the following direction: in all experimentation use the one-sided test which is powerful for the alternative in the direction of the sample mean, $\overline{X}$. However, to shift back and forth from one one-sided test to another, depending on the direction of $\overline{X}$, makes what would nominally seem a one-sided test at the .05 level, actually a two-sided test at the .10 level. To follow this procedure, and speak of the .05 level of significance, creates the impression that the degree of significance found is higher than it really is.

It should be pointed out that the typical one- and two-sided tests are only special cases from the hierarchy of statistical tests that can be devised. Although tests can be based on a large variety of critical regions, we are ordinarily most concerned with hypotheses operationally parallel to the one- and two-sided cases, and more frequently the latter than the former.

Where can we justify one-sided tests for one and two sample inferences in psychology? Such tests seem seldom appropriate except for certain limited practical problems. It is conceivable that in comparing consumer preference for two products one might desire a test which was most powerful for the alternative—"preference for the proposed product A is significantly greater than preference for the present product B"—and where it was of no concern to be able to establish, if true, that preference for B was significantly greater than for A. Except for similar examples, it seems that in most psychological research it is of definite interest to have a test which is powerful in terms of being able to detect a difference opposite in sign from the predicted one. Even in studies of teaching methods, for example, where we would not adopt a new procedure unless average results exceeded a given value, it would often be pertinent to know if the new method were significantly poorer than the established one.

---

[1] The writer is indebted to Professor Wilfrid J. Dixon for clarifying in a personal communication how the logic of the one-sided two-sample test parallels that of the one-sided test of an exact hypothesis.

The rationale of one- and two-sided tests has been discussed in the light of a number of questions which typically confront the experimenter. In considering problems like this, we should remember that in terms of action we base on the outcome of statistical tests, the difference between a one-sided and a two-sided test may be important in one case and negligible in another. At the same time, analysis of such differences in statistical procedure, even when minor in terms of a particular conclusion, may be of assistance in clarifying processes involved in experimental reasoning.

## REFERENCES

1. BURKE, C. J. A brief note on one-tailed tests. *Psychol. Bull.*, 1953, **50**, 384-387.
2. DIXON, W. J., & MASSEY, F. J. Introduction to Statistical Analysis. New York: McGraw-Hill, 1951.
3. HICK, W. E. A note on one-tailed and two-tailed tests. *Psychol. Rev.*, 1952, **59**, 316-318.
4. JONES, L. V. Tests of hypotheses: one-sided vs. two-sided alternatives. *Psychol. Bull.*, 1952, **49**, 43-46.
5. KENDALL, M. G. The Advanced Theory of Statistics, Vol. 2. (2nd Ed.) London: Griffin, 1948.
6. MARKS, M. R. Two kinds of experiment distinguished in terms of statistical operations. *Psychol. Rev.*, 1951, **58**, 179-184.
7. ———. One- and two-tailed tests. *Psychol. Rev.*, 1953, **60**, 207-208.
8. MOOD, A. M. Introduction to the Theory of Statistics. New York: McGraw-Hill, 1950.

*Department of Psychology*
*University of Denver*
*Denver 10, Colorado*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 103-126.

# THE INTERGROUP RELATIONS WORKSHOP: ITS PSYCHOLOGICAL AIMS AND EFFECTS*

*Department of Social Relations, Harvard University*

Daniel J. Levinson[1]

The past decade has seen a remarkable proliferation of attempts to apply psycho-social theory in the solution of practical social problems, particularly in the field of intergroup relations (e.g. 2, 3, 4, 8, 11). However, the application of "human engineering" methods has far outstripped our ability, and even our readiness, to assess the effectiveness of these methods in fulfilling their intended functions. This paper is a preliminary attempt to evaluate one such method. More specifically, I should like to formulate the major aims of a Workshop on Intergroup Relations given at Harvard University in the 1951 Summer Session and to present empirical evidence concerning the degree to which these aims were realized. The report thus has a twofold purpose: a *practical-technological* one of assessing the value of an educational method; and a *theoretical* one of understanding the processes by which a particular situation produces various psychological effects.

## A. The Harvard Workshop

The Harvard Workshop met five mornings a week for an hour of lecture and discussion; in addition, the group split up into teams which worked on special projects and met separately one afternoon a week. The group as a whole went on several picnics, and extra-curricular activities were encouraged but not organized "from above." The entire session lasted six weeks. It was less intensive than certain workshops in which the members live together in a dormitory and spend most of their waking hours in joint educational-recreational activities; however, many workshops involve less intensive work and interaction. The workshop was taught by the writer in collaboration with Miss Carol M. Smith, a high school teacher and dean. Except for a few lectures by distinguished visitors, all the teaching was done by the two course instructors.

---

*Received in the Editorial Office on February 19, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] I wish to acknowledge my indebtedness to Carol W. Smith, who collaborated with me in giving the Harvard Workshop; also to Lora W. Heims and Paul E. Sapir for their aid in various phases of the research. The research was facilitated by a grant from the Laboratory of Social Relations, Harvard University, and from the Lt. Hirschel L. Gutman Foundation for the Advancement of American Ideals, Inc

The size and social composition of the Harvard Workshop is probably fairly typical. There were 28 members. Slightly more than half of them were residents of Massachusetts; the remainder came from other regions of this country, except for one foreign exchange student from Germany. Occupationally, 24 were teachers in public, private and parochial schools; the remaining four included a policeman, a school superintendent, the exchange student, and a social agency worker. There were 5 Negroes and 23 Caucasians. By religion, there were 9 Catholics, 7 Jews, 11 Protestants and one unaffiliated. The workshop members varied in age from approximately 20 to 60 years, most of them being in their thirties.

## B. INTENDED FUNCTIONS OF THE WORKSHOP

We are in need of more explicit, systematic consideration of the intended functions of this and other educational methods. By formulating the *intended* effects of the workshop, we provide ourselves with relevant criteria for the evaluation of its *actual* effects. Concern with both intended and actual effects may stimulate continued reformulation of the ideal functions of the workshop and, at the same time, of our underlying philosophy of education. Our aims at a given time may be misplaced, unimaginative, impractical, overly practical, rigidly moralistic, chaotically amoral. In short, by becoming more aware of what we now want to achieve, we can better evaluate our actual achievements and, equally important, we can constructively criticize and modify our present aims.

The present Workshop, like most others in intergroup relations, had the following intended functions:

1. *To provide information concerning various aspects of group life and skills for dealing with everyday "intergroup" problems.* Within this broad area numerous problems are available for emphasis. The four main substantive fields covered in the Harvard Workshop were these: psychological aspects of group conflict (with focus on the rôle of the dominant group); effects of discrimination and reactions to it within various subordinate groups; sociological aspects of group conflict (mass media, industry, housing, education, religion, and the like); methods for inducing change toward democratic functioning in individuals, primary groups, or larger institutional structures. The only assigned common reading was the recently published anthology edited by Rose (11).

It is often assumed that the workshop members have a clear conception of what prejudice is, that they have at least the general outlines of a democratic ideology of group relations, and that the workshop is primarily a means of implementing their democratic convictions with an armament of information and technical skills. Indeed, workshop leaders who place especial emphasis on the knowledge-skill function of the workshop sometimes make this assumption implicitly and without question. I should like to question it, however, and to suggest a second major function of the workshop:

2. *To help the members clarify the meaning of concepts such as "prejudice," "discrimination," "democratic group functioning" and the like; to further their growth toward democratic modes of thinking and relating to self and others.* This is perhaps the most difficult aim to realize. It is more pleasant for all of us to focus on what we regard as evil or harmful in others than to deal with the same tendencies in ourselves. It is harder to achieve self-insights than to acquire facts. In fulfilling this function, the instructor must find some compromise, never entirely adequate, between pedagogy and group therapy. He takes the risk of activating formidable anxieties but must prevent intrapersonal and interpersonal disruptions. And he has more difficulty than ever in expressing conviction without dogmatism, permissiveness without indifference. This aim was given a high priority in the Harvard Workshop. There was much discussion of general principles of democratic *vs.* autocratic ideology and group structure. The discussion extended beyond the sphere of ethnic group relations to problems of international relations, religion, the family, industry and the like. We were led to a consideration of psychological factors in the formation of ideology, and to the nature of authoritarian and equalitarian personality syndromes.

3. *To provide an intense, supportive living experience.* This is a less tangible function and one whose effects are not easily measured, yet its importance is considerable. The workshop members are highly diversified with regard to various group characteristics such as religion, skin color, ethnic and regional background, and sex. (The chief characteristics which merit further diversification are occupation and class, for there is a conspicuous concentration of middle class teachers and social service personnel.) Many workshoppers belong to groups which have a subordinate status in their home communities. Almost all of them live and work in a milieu that regards "liberal" ideas with suspicion, that doubts or actively opposes the principles espoused in the workshop. The workshop experience can, therefore, provide tremendous emotional support for wishes and values that go against the grain of the ordinary social environment and are continually undermined by it. For those workshoppers who belong to highly oppressed minority groups, the workshop often represents a kind of utopian experience, an opportunity for joint participation and individualized relationship unimpeded by

stereotypes and artifices. And those who belong to no major subordinate group sometimes derive great psychological value from the shared experience with individuals who are ordinarily inaccessible because of membership in a socially distant group.

There were wide individual differences in the ability to enjoy and to gain from this aspect of the workshop situation. For some workshop members the sense of dichotomy between various groups, and the tendency to experience self and others in group-label rather than individualized terms, were so strong that the workshop failed entirely in this regard (and in most others as well). Several others were able from the start to approach other persons as individuals, with minimal concern for ethnic, religious, or other group memberships. The great majority of the group fell between these two extremes. Although their values were in the direction of "democracy" and "tolerance," a life-situation that brought them into close contact with members of other groups brought to the surface all sorts of ambivalences and irrational expectations. Perhaps the term *fascination* best connotes the complex mixture of positive and negative feelings involved. Thus, for some of the white students the thought of a close relationship with a Negro had a kind of illicit quality: it seemed attractive; it offended no conscious values; and yet there was something remarkable and almost fearful about it. Analogues of this could be found in some of the Negro students, and in other white students with respect to the Catholic-Jewish-Protestant distinction.

## C. Procedures Used in the Study of Change

An ideal evaluation study should assess the degree of actual change in *all* the modes of thinking, feeling, and acting that the workshop *attempted* to change. These changes should be assessed immediately following the workshop and again after an interval of several months or longer. Unfortunately, it is difficult to achieve an adequately comprehensive and rigorous evaluation study.

The present evaluation research was chiefly concerned with the second major workshop function discussed above. Our chief question was: *To what degree did the workshop members become more democratic and understanding in their approach to self, other persons, and social problems generally?* In principle, the most desirable way of studying such changes is by means of systematic behavioral observation and clinical study. However, a study of this type was not feasible. Instead, the primary technique used for the measurement of change was the following set of six opinion-attitude scales. Three of these scales (*E*, *F*, and *PEC*) I derived from forms initially developed by several colleagues and myself in a study of authoritarian personality (1). The remaining three, of more recent vintage, are outgrowths of the earlier research. Each scale contains 12 items.

(*a*). The Authoritarianism (*F*) Scale contains a variety of ideas concerning authority, work, sex, aggression, human nature, and the like. High scores presumably indicate personality trends such as punitiveness, authoritarian submission, stereotypy, and projectivity.

(*b*). The Ethnocentrism (*E*) Scale deals with intergroup relations within this country and at the international level. High scores presumably indicate a generalized tendency to reject various outgroups such as Negroes, Jews, foreigners and other nations, and to idealize the corresponding ingroups. Low scores are associated with rejection of both outgroup prejudice and ingroup idealization.

(*c*). The *Religious Conventionalism* (*RC*) Scale contains a variety of ideas about the church, the Bible, God, prayer and the like (7). High scores reflect a conception of God as punishing power figure and of the church as absolute moral authority, as well as a marked emphasis on faith, tradition, and conformity to institutional forms. Low scores, on the other hand, represent a religious humanistic or a nontheistic approach which emphasizes reason, personally derived values, and a naturalistic rather than supernaturalistic view of the world.

(*d*). The *Politico-Economic Conservatism* (*PEC*) Scale provides a measure of general liberalism (low scores) *vs.* conservatism (high scores) concerning such issues as economic security, taxation, and the government-business-labor balance of power.

(*e*). The *Traditional Family Ideology* (*TFI*) Scale deals with husband-wife relations, parent-child relations, conceptions of authority, masculinity and femininity, and so on (6). High scores indicate strong emphasis on discipline in child-training, on the dominant-assertive male and the rigidly conventional female. Low scores indicate less dichotomous definition of sex rôles and greater emphasis on self-expression.

(*f*). The *Internationalism-Nationalism* (*IN*) Scale, a recent extension of the Ethnocentrism Scale, is concerned with current international issues and with general values regarding leadership, civil liberties and national policy (5). The *IN* and *E* scales correlate highly (.7 to .8) with each other and nationalism as defined here is essentially an expression of the ethnocentric orientation at the international level.

These six scales, taken together, may be said to comprise an "Authoritarianism" battery. Previous studies indicate that the *F* Scale correlates in the neighborhood of .6-.7 with *E, IN,* and *TFI.* All four of these scales yield correlations averaging about .5 with *RC* and .3-.4 with *PEC.* Since the workshop was concerned with the development of a democratic approach in the broadest sense, and not merely with specific group problems, these scales seemed appropriate for evaluation purposes.

The scales were presented, together with a number of personal-social data questions, in a questionnaire known as the *TAP* Social Attitude Battery, Form *FERPTI*. Copies of the questionnaire are available to interested researchers. The questionnaire was administered during the second workshop meeting prior to any expression of viewpoint by the instructors. It was given again a few days before the end of the six-week session. On both occasions it was presented as part of a larger research project, and subjects were asked not to sign their questionnaires. (Initial and final questionnaires were paired by a comparison of birthdates and other personal-social data.)

## D. INITIAL PERSONAL-ATTITUDINAL CHARACTERISTICS OF THE WORKSHOP GROUP

Let us consider first the scale data for the Harvard Workshop group at the start of the workshop, in relation to data on three comparison groups. The procedure used in scoring the scale responses has been described in detail elsewhere (1). A brief summary will suffice here. The subject is asked to indicate the degree of his agreement (+1, +2, +3) or of his disagreement (—1, —2, —3) with each scale item. Items are scored on a 1-7 scale so that higher scores indicate a more authoritarian position. The item total is divided by the number of items to put us back on a 1-7 series. Scale scores of over about 5 points are relatively high, while scores falling between 1 and 3 are relatively low and indicate an orientation opposite to that given by the scale name.

For comparative purposes we have data on three groups in addition to the Harvard Workshop. (*a*) The *Control Group* is an Education class in the Harvard Summer Session which was given "initial" and "final" administrations of the questionnaire at the same time as the Workshop group. It is similar to the workshop group in age, occupation and several other characteristics. (*b*) The *Columbia Workshop* on Intergroup Relations resembles the Harvard one in size and psycho-social composition. (*c*) The *College Students*, 236 subjects, comprise almost the entire sophomore class in General Education at a large university in the Boston area. On the basis of previous evidence (1) it can be assumed that the scores earned by this group will be slightly lower—more "democratic"—than the average for a more representative urban community sample.

The Mean, Standard Deviation and Range of the scores made by each of the four groups on the various scales are presented in Table 1. On the Ethnocentrism (*E*) Scale the Mean of 2.3 for the Harvard Workshop group is near the "democratic" (low) extreme and well below that of 2.9 for the College Students, the difference being statistically significant at well beyond the 5 per cent level. This result is hardly surprising. Most workshoppers feel strongly about the injustice of discrimination and are eager to do something about it. On the other scales, however, the Harvard group shows greater variability and less concentration of scores on the democratic side of the continuum. Thus, the Harvard Mean is 3.3 on the Authoritarianism (*F*) Scale, a value only slightly below that of 3.5 for the College Students, and slightly above the Control Group's 3.1.

TABLE 1

INITIAL STANDING OF THE HARVARD WORKSHOP GROUP, AND OF SEVERAL COMPARISON GROUPS, ON THE VARIOUS SCALES

| | Group | | | |
|---|---|---|---|---|
| Scale data<br>Number of cases | Harvard Workshop*<br>28 | Control Group<br>19 | Columbia Workshop<br>28 | College Students<br>236 |
| F Scale: | | | | |
| Mean | 3.27 | 3.07 | 2.90 | 3.49 |
| *SD* | 1.16 | 0.89 | 1.02 | 0.84 |
| Range | 1.5-6.2 | 1.6-4.6 | 1.2-5.1 | 1.0-6.2 |
| E Scale: | | | | |
| Mean | 2.30 | 2.73 | 2.30 | 2.95 |
| *SD* | 0.88 | 0.94 | 0.89 | 0.90 |
| Range | 1.0-4.6 | 1.0-4.7 | 1.0-4.8 | 1.2-5.7 |
| RC Scale: | | | | |
| Mean | 4.45 | 3.22 | 4.37 | 4.06 |
| *SD* | 1.66 | 1.51 | 1.61 | 1.55 |
| Range | 1.0-6.8 | 1.0-6.5 | 1.0-6.6 | 1.0-7.0 |
| PEC Scale: | | | | |
| Mean | 3.61 | 4.01 | 3.43 | 4.14 |
| *SD* | 1.16 | 1.10 | 1.08 | 0.97 |
| Range | 1.1-5.8 | 1.7-6.1 | 1.3-5.1 | 1.3-6.6 |
| TFI Scale: | | | | |
| Mean | 3.07 | 2.63 | 2.57 | 3.40 |
| *SD* | 1.21 | 1.01 | 0.91 | 0.96 |
| Range | 1.0-5.2 | 1.0-4.4 | 1.2-4.2 | 1.2-5.6 |
| IN Scale: | | | | |
| Mean | 3.25 | 3.23 | ** | ** |
| *SD* | 1.40 | 1.32 | ** | ** |
| Range | 1.0-6.4 | 1.0-5.3 | | |

*The Harvard Workshop differs significantly (5 per cent level) from the Control Group only in the case of the *RC* Scale. On no scale do the Harvard and Columbia Workshops differ significantly. The Harvard Workshop and the College Students differ significantly on the *E* and *PEC* Scales.

**The *IN* Scale was not included in the questionnaire given to these groups.

Moreover, there is a striking and statistically significant discrepancy of one full point between the *E* Mean (2.3) and the *F* Mean (3.3) of the Workshop group. A similar *E-F* difference was found in a previous workshop study (8). The difference in the case of the present College Student sample is only .5, and an examination of data from a variety of samples reported earlier (1) shows that the *E-F* difference seldom reaches .3. Similar discrepancies are found between the *E* and *TFI* Scales. On the Religious Conventionalism Scale, the Harvard Workshop Mean (4.45) is well above that of the College Students (4.1) and significantly greater than the Control Group Mean (3.2). Politically, the Harvard group is slightly more liberal than the two non-workshop groups, but its *PEC* Scale Mean of 3.6 and its wide range of individual scores (1.1-5.8) do not suggest any special leanings in a liberal direction.

In short, though most of the Harvard workshoppers have strong democratic convictions with regard to minority problems, these convictions are often weak or absent when other social issues are involved. A sizable minority of the group earns *F* and *TFI* scores which are clearly on the high ("undemocratic") side of the continuum; and the total group mean on both of these scales is only slightly lower than those of the unselected College sophomores. We are thus led to believe that, for some of the workshoppers, the opposition to prejudice is *not* imbedded in a larger psychological context of democratic ideology and equalitarian personality. These are perhaps the representatives of "tolerance" in the negative sense of the term: persons who "help the unfortunates" but are not able to identify with them, to comprehend the responsibility of the *status quo* for the plight of a subordinate group, or to conceive of a social system in which group subordination, at least in its more gross forms, is eliminated. This interpretation is supported by numerous observations made during the course of the workshop.

Some additional insights into this problem are provided by workshop data on various *E, IN,* and *F* Scale items. The ethnocentric statements which are most actively rejected by the group, and on which they accordingly earn the lowest group Mean, are those (six) referring to Jews and Negroes. However, on other *E* items which express more general and subtle forms of discrimination or which idealize important membership groups (e.g., America, one's own family), there is less uniform disagreement and a higher group Mean. Again, the Workshop Mean on the *IN* (Nationalism) Scale is 3.25, a value which approximates the *F* Mean and indicates that some workshop members tend to think of international relations in moralistic ingroup-outgroup terms, to think of foreign policy chiefly in terms of American military power, and to exhibit other aspects of an ethnocentric orientation with respect to international relations.[2] Finally, the Means of the 12 *F* Scale items vary considerably. The most rejected (low Mean) items are those which most clearly express destructive, punitive, vindictive tendencies. There is, however, a higher incidence of agreement with ideas involving conformism, submission to conventional authority, and emphasis on hard work and "keeping busy" at the expense of inner life and self-expression.

These findings, again seen in relation to direct workshop observations, suggest that there are a number of relatively authoritarian workshop members. These individuals represent primarily one variant of the general authoritarian syndrome. Repulsed by destructiveness, cynicism, and the grosser forms of social injustice, they are dedicated to ideals of kindness and social harmony. They would not *knowingly* engage in actions that cause suffering or hardship, and the expression of *recognizable* prejudice by their associates disturbs them. They are authoritarian, however, in their inability to accept deviance and in the imperiousness of their inner requirement to be on good terms with conventional authority—whether that authority be a parent, a community leader, an institutional pressure, or merely the force of prevailing opinion. They are estranged from their own inner lives, from those wishes and fantasies that threaten their harmonious relation to an authority they must conceive of as benevolent. For the same reasons, they are not ready to recognize the degree to which discrimination is maintained or furthered by conventional friends, authorities, and institutions. Moreover, members of subordinate groups are in a sense as threatening as the authority-criticizing ideas and wishes within these subjects. For "minority" individuals, if known well, would provide an unconscious reminder of the dangers that can befall the deviant.

The foregoing discussion has centered on the Harvard Workshop group. We may note in passing that the initial standing of the Control Group is very similar to that of the Workshop on the *F* and *IN* Scales; the Control Group is nonsignificantly higher on the *E* and *PEC* Scales, and slightly lower than the Workshop on *TFI*. There is a statistically significant difference between the two groups only in the case of the *RC* Scale, on which the Workshop earns an unusually high Mean.

---

[2] Another expression of the same tendencies is to be found in the literature of some "brotherhood" organizations, viz., the idea that we must avoid inter-religious and inter-racial conflict in America in order to be stronger in the struggle with Russia and our "national enemies." Proponents of this view are usually more concerned with achieving a sense of existing national harmony than with working actively to eliminate the inequities and discriminatory practices that now exist. They are suggesting, in effect, that we forget our pressing intergroup problems within the nation and instead focus our hostilities on extranational targets.

The Columbia Workshop shows a profile of scale means very similar to that of the Harvard Workshop. Both are conspicuous in relation to the other groups for their low degree of ethnocentrism and their high degree of religious conventionalism. However, the Columbia group tends slightly but consistently to be more liberal and nonauthoritarian than the Harvard group. We may also surmise—and the group difference in maximum score on the *F* and *TFI* Scales supports this—that there were fewer Columbia than Harvard workshoppers of the authoritarian pattern just discussed.

I have been assuming thus far that the characteristics measured by the various scales go together to form a syndrome in the Harvard Workshop group just as they have in previous groups (see Section *C*, above). Evidence concerning the adequacy of this assumption is contained in Table 2, which gives the rank-order correlations among the six scales for the Harvard Workshop. In general, the correlations are of the expected order of magnitude, especially in view of the decreased dispersion of scores on the *E* and *TFI* Scales. Thus, the correlation of .77 between *E* and *IN* supports the thesis that nationalism is one aspect, albeit a separable one, of general ethnocentrism. The remaining correlations among the *E, IN, F,* and *TFI* Scales vary between .62 and .72, with the sole exception of the *IN-F* value of .53. All of these scales except *F* correlate about .6 with the Religious Conventionalism Scale. Their correlations with Politico-Economic Conservatism are considerably lower (.26 to .45), though only slightly below the level obtained these days with more diverse groups. The diminished *PEC* Scale correlations suggest that the workshoppers' political views are poorly crystallized and not geared in with other views. This may be a consequence of political naiveté or disinterest specific to the workshop group. I suggest, however, that it is due in part to a more general *apoliticalization* of approach to social problems which

TABLE 2

RANK-ORDER CORRELATIONS AMONG THE VARIOUS SCALES* IN THE INITIAL TESTING OF THE HARVARD WORKSHOP

| | E | RC | PEC | TFI | IN |
|---|---|---|---|---|---|
| F | .62 | .35 | .26 | .70 | .53 |
| E | | .64 | .33 | .72 | .77 |
| RC | | | .28 | .64 | .60 |
| PEC | | | | .45 | .36 |
| TFI | | | | | .65 |

*For a brief description of each scale, see Section *C*.

has been increasing on the national level in recent years. This trend may account for the fact that the *PEC* Scale correlations have been steadily decreasing since 1947.

In summary, the workshop contains a relatively authoritarian subgroup which tends also to be relatively the most prejudiced and nationalistic, and the most autocratically tradition-bound in its orientation toward family, religion, and politics. In addition, there are several members who represent a contrasting syndrome of equalitarian personality trends and a more liberal, humanistic approach to social problems. Finally, there is a "middle" group about which we know all too little.

## E. Psychological Changes between Beginning and End of the Workshop

In comparing initial and final scale scores, we are concerned first with the question of *individual stability:* to what degree do individuals retain the same rank-position in the group, from initial to final testing, on each scale? An answer to this question is given by the rank-order correlations between initial scores and final scores on each scale. For the six scales, *F, E, RC, PEC, TFI,* and *IN,* the test-retest correlations are .87, .88, .89, .69, .93, and .90, respectively. Regarded as reliabilities, these correlations are, with the exception of that for the *PEC* Scale, quite satisfactory for most purposes; and all of them approximate the reliability values obtained on other groups by the split-half (Spearman-Brown) method. Looked at from the point of view of stability of individual response, these correlations indicate that there were relatively few gross fluctuations in score from initial to final testing. This does not mean that no shifts in score occurred; it does mean that whatever change occurred was well distributed throughout the group, so that the final rank-order of our subjects on each scale was very similar to that which obtained at the start of the workshop.

Our second question concerns over-all *group change.* To the extent that the intended personal-ideological functions of the workshop were fulfilled, we expect a significant downward shift in the group mean on each scale, that is, a change in the direction of a more equalitarian orientation toward self and society. In more formal terms, we predict that the Harvard Workshop will show a significant downward change on all scales, and that the Control Group will show no significant change. To the extent that its aims and methods were similar, the Columbia Workshop should also drop significantly.

The pertinent data are presented in Table 3. The general pattern of results can be stated simply: *The Harvard Workshop has undergone a signifi-*

*cant change in the predicted (equalitarian) direction on five of the six scales.* On the sixth scale, *PEC,* the shift is also downward but to a degree not significant at the minimally acceptable (5 per cent) level of confidence.

TABLE 3

CHANGE IN AVERAGE SCORE ON THE VARIOUS SCALES BETWEEN INITIAL AND FINAL TESTINGS (3 SUMMER SESSION CLASSES)

| Scale data<br>Number of cases | Harvard Workshop<br>28 | Control Group<br>19 | Columbia Workshop<br>28 |
|---|---|---|---|
| F Scale: | | | |
| Final Mean | 2.87 | 3.24 | 2.69 |
| Change ($M_2$—$M_1$) | — .40 | + .17 | — .21 |
| “t” | 3.2** | 1.4 | 1.6 |
| E Scale: | | | |
| Final Mean | 2.12 | 2.71 | 1.94 |
| Change ($M_2$—$M_1$) | — .18 | — .02 | — .36 |
| “t” | 1.8* | 0.2 | 3.3** |
| RC Scale: | | | |
| Final Mean | 3.82 | 3.11 | 4.17 |
| Change ($M_2$—$M_1$) | — .63 | — .11 | — .20 |
| “t” | 3.5** | 0.7 | 1.1 |
| PEC Scale: | | | |
| Final Mean | 3.44 | 3.80 | 3.34 |
| Change ($M_2$—$M_1$) | — .17 | — .21 | — .09 |
| “t” | 1.2 | 2.0* | 0.8 |
| TFI Scale: | | | |
| Final Mean | 2.66 | 2.76 | 2.30 |
| Change ($M_2$—$M_1$) | — .41 | + .13 | — .27 |
| “t” | 3.4** | 1.4 | 2.4* |
| IN Scale: | | | |
| Final Mean | 2.76 | 3.23 | —— |
| Change ($M_2$—$M_1$) | — .49 | 0.00 | —— |
| “t” | 4.0** | 0.0 | —— |

In the table, “$M_2$” is the final mean, “$M_1$”, the initial mean; since change is computed as $M_2$—$M_1$, negative values indicate a drop in mean, positive values, a rise.

A *“t”* value followed by a single asterisk is significant at the 5 per cent level of confidence; a double asterisk indicates the 1 per cent level or better. For a sample of 28 cases, using a one-tailed test of significance (since the workshop predictions had directionality), a *“t”* of at least 1.70 is needed for the 5 per cent level, and 2.47 for the 1 per cent level.

The Columbia Group did not receive the *IN* Scale.

On no scale is there an upward shift. *The Control Group, on the other hand, shows no consistent change in either direction:* On two scales (*E* and *IN*) there is no change; on another two (*F* and *TFI*) there is a slight but not significant shift in an upward (authoritarian) direction; the *RC* and *PEC* Scales show a downward shift but only the latter reaches the 5 per cent level of significance. The over-all Control Group change is perhaps best viewed as falling within the limits of chance variation. *The Columbia Workshop, like the Harvard Workshop, exhibits a consistent downward shift on all scales; but the change attains statistical significance only for the E and TFI Scales.* This group might well have shown greater change on measures pertaining more directly to the issues covered by it.

The change in the Harvard Workshop would seem to be due in part to the explicit discussion of authoritarian and equalitarian orientations and their manifold expressions in various aspects of social life. It should be noted, however, that the Harvard Workshop did not discuss the various scales as such, nor even the specific content of the scale items. The actual procedure was as follows: the students were presented didactically with certain general ideas concerning authoritarianism and equalitarianism as personality syndromes, and concerning democratic *versus* autocratic forms of group process and social structure. This was followed by group discussion—often heated, sometimes quite tense, and always exciting—of the application of these ideas to specific situations or problems: of the classroom, of parenthood, of intergroup conflict, of religious and economic functioning. Many of the questions raised by the group are well-nigh universal in workshops and other intellectual forums: How permissive or controlling should the teacher or parent or other authority be? What should be left to individual conscience, what to external institutional control? What does it mean to be "unprejudiced"? What are the values and disadvantages of self-insight?

Though most members could agree on the nature of democratic principles in the abstract, there was much difference of opinion as to their application in specific situations. In addition, there were some group members who disagreed with the prevailing view on basic principles, and there were a few who never quite got the points at issue. The instructors attempted to formulate the general ideas as a basis for discussion. However—and this is crucial for the evaluation of the scale findings—we did not try to formulate a detailed set of slogans as to "what to do" and "what not to do" in particular situations. Rather, we presented a general approach to be accepted or rejected as an individual matter and, if accepted, to be spelled out and applied by each individual within the framework of his personal experience. In short, the workshop content and the scale content were by no means identical. The final administration of the questionnaire was not, so to say, like a final examination in which the student reads each question (scale item), asks himself what the instructor defines as the "correct" answer, and responds accordingly. It may well be that part of the drop in scale scores is due to an attempt, by some workshoppers, to meet the sensed expectations of the instructors. I propose, however, that the change shown on the scales reflects as well a genuine increase in the strength of democratic convictions. This point will shortly be elaborated.

Thus far we have considered only the over-all changes found in the Harvard Workshop as a whole. Our final research question concerns the differential effects of the workshop on various *psychological subgroups* within it. Do the participants' personality characteristics at the start of the workshop influence how much they learn from the workshop experience? This phase of the research brings us one step closer to a psychological understanding of the workshop process.

The best personality criterion available within the setting of the present research is the *F* Scale, our crude index of authoritarian personality. It permits us to specify the question about the rôle of personality in learning, as follows: How do the Workshop members who are initially the most authoritarian (High Third on *F*), the most equalitarian (Low Third), and "intermediate," differ in the degree to which their views change over the course of the workshop? We have accordingly determined the average change undergone by each of these subgroups on the various scales.

When we attempt to predict the outcome in this case, three main possibilities present themselves. (*a*) The *High* Third will change the most: The Highs on *F*, being also relatively the highest scorers on the other scales (as indicated by the correlations in Table 2), have the most room for change. Since the workshop tends to disvalue the authoritarian orientation, the Highs may respond in a superficially more democratic manner out of conformism even when their inner convictions remain unchanged. Finally, the tendency of "regression to the mean," which is inevitably present on the second administration of any imperfect measuring instrument, operates to produce a spurious downward shift in the High Group. (*b*) The *Middle* Third will change the most: Its scores are, after all, high enough to provide some room for change. The middle position, being a less structured one, may be more amenable to change. Since the middle group in the present sample is slightly on the equalitarian side to begin with—its initial *F* scores fall between 2.6 and 3.7—it may be ready to form more consistent and intense equalitarian convictions. On the other hand, holding a "neutral" position on controversial issues may have essential emotional functions for these persons; perhaps it is so important to them to "see both sides" of every controversy that they can never be induced to go out on a limb. (*c*) The *Low* Third will change the most: It should be more receptive than the other two to the dominant workshop orientation. However, it has the least room for downward change, and the "regression to the mean" tendency operates in the direction of higher scores. Prediction for the Low group is accordingly a difficult matter.

There is an additional set of factors that help determine which third will change the most, namely, the external pressure in the workshop situation. Different workshop contexts may well have differential effects on the High, Middle, and Low $F$ groupings; e.g., one workshop may be particularly effective with Middles and another with Highs. This problem can be discussed more fruitfully after the findings have been presented.

Table 4 shows the average change in each $F$ scale grouping on the various scales used. The results are consistent and clear-cut: *The workshop had its maximum effect on those who were initially Middle in degree of authoritarianism, and its minimal effect on those who were most authoritarian.* The average downward shift for the Middle Third was .54, three times greater than that of .18 for the High Third. The change for the Middles was at least twice that of the Highs on every scale except *PEC* (the perennial exception). Considering its low initial standing, the Low Third underwent a remarkable change (.41) which nearly equalled that of the Middles. And the Lows' change was exceeded by the Highs' on only one scale (*PEC*). In short, the students' initial personality characteristics did in fact influence their receptivity to new ideas. The more authoritarian students showed only a negligible change; those who were initially equalitarian became more consistent and more intense in their convictions; and the "middles" made a noticeable step toward a more democratic orientation.

TABLE 4

DEGREE OF CHANGE ON EACH OF THE SCALES AS A FUNCTION OF INITIAL STANDING ON THE $F$ SCALE (HARVARD WORKSHOP)

| Initial standing on $F$ scale | Degree of change ($M_2$—$M_1$) | | | | | | |
|---|---|---|---|---|---|---|---|
| | F | E | RC | PEC | TFI | IN | *Over-all* |
| Low Third ($N = 9$) | —.29 | —.27 | —.92 | —.14 | —.33 | —.50 | —.41 |
| Middle Third ($N = 10$) | —.65 | —.23 | —.88 | —.01 | —.84 | —.64 | —.54 |
| High Third ($N = 9$) | —.29 | —.05 | —.06 | —.37 | —.01 | —.32 | —.18 |
| Total group ($N = 28$) | —.40 | —.18 | —.63 | —.17 | —.41 | —.49 | —.38 |

Before attempting any further interpretation of these findings, let me note briefly the corresponding data for the Columbia Workshop. High, Middle, and Low groups were again derived on the basis of initial $F$ score. The average downward shifts shown by these subgroups, on the total battery of scales, are .35, .22, and .11, respectively. The over-all sample change is .22. The four corresponding values for the Harvard Workshop (see the extreme right column of Table 4) are .18, .54, .41, and .38. Two differences stand out immediately: First, that the Columbia group as a whole changed less; and second, that the pattern of change in the Low, Middle, and High Thirds is quite different in the two workshops. In the Columbia Workshop, change is an inverse function of initial *F* score. The Highs change the most, the Lows the least.

It would appear, then, that there is no simple answer to our original question: Which "authoritarianism" grouping is likely to be most influenced by a workshop? It depends on the workshop. At this point we have, in a sense, merely demonstrated the ancient psychological adage which Lewin has given a modern formulation. To paraphrase Lewin: Learning is a function of the *person* and the *situation.* In any given situation (workshop), individuals will learn differentially, depending on their personalities. On the other hand, individuals having similar personalities will learn differentially, depending on the psychological character of the situation.

In order to say with confidence why the two workshops in this study should have produced such different effects, we need to know more about the psychological characteristics of their members, and about the organization of the workshops—the teaching methods, the teacher personalities, the ideas emphasized, the kinds of pressures and opportunities presented. In the absence of direct evidence, I should like to suggest a few, intuitively derived hypotheses about the interaction of situational and personality factors.

1. The authoritarian-equalitarian continuum of personality is an important one for workshop leaders to take into account in planning the workshop. No one approach will be equally effective with the Highs, Middles, and Lows on *F,* so that a variety of approaches must be attempted.

2. Nevertheless, I doubt that any workshop, no matter how multifaceted its approach, can be highly effective with all of the *F* Scale Thirds. (I refer to "*F* Scale Thirds" merely to keep continuity of terminology; Thirds are not necessarily the best basis for forming subgroups, nor is the *F* scale necessarily the best criterion.) For, even with the most varied approach, every workshop is limited by its own special operating principles, its dominant orientation and atmosphere. In addition, the workshop leaders necessarily form relationships of different types and intensities with the various students; these relationships, which are influenced to some degree by similarities in degree of authoritarianism, in turn partly determine how much individual workshoppers learn. Those participants who are alienated or left indifferent will gain less from the workshop experience.

The following two hypotheses concerning the present study are derived from those just presented. Since they are *ad hoc,* they require testing by means of new, appropriately designed research in which the predictions are stated beforehand.

3. The Harvard Workshop had the most effect on Middles and Lows, and the least effect on Highs, because it was in fact slanted most toward the former subgroups and because it tended somewhat to threaten and alienate the Highs. The workshop emphasized the psychological aspects of group conflict. The participants were encouraged to examine their values, their readiness to modify institutional patterns and to relate freely with other individuals regardless of group memberships, their opinions on such "private" matters as religion and family relationships. A great premium was placed on self-insight, especially with regard to the subtle ways in which our interpersonal experiences are fogged and distorted by irrational anxieties and stereotypes. This is not to say that discrimination is a purely individual matter; the point emphasized was that irrationalities within the individual limit both his capacity to relate freely with others and his ability to support those programs of social action that may reduce discrimination at the institutional level.

This workshop orientation found its greatest appeal among the Low and Middle authoritarianism groupings. It led them not only to look within themselves but also to relate more intensely with each other. Frequently relationships that began, on both sides, with unspoken questions such as, "What can I learn about Negroes?" (or whites or Christians or whatever the group-label involved) or, "Can I really be friends with one?" came to be genuine inter-*person* friendships in which group labels were relatively unimportant.

However, the more authoritarian participants were threatened by this psychological setting. It was too intense; it sought to open too many inner doors; it criticized personal characteristics seemingly unconnected with "minority problems"; and it presented too few techniques and ready-made solutions for dealing with these problems. The alienation of these persons from the rest of the group might have been avoided had there been an additional staff member of a more "middle" variety with whom they could more readily identify. Also, more could have been done to help them assimilate the ideas presented and to reassure them as to their acceptance by the staff and the other participants.

4. Not having participated in the Columbia Workshop, I can offer only a few highly tentative hypotheses about it. This workshop had a larger staff

and more guest speakers than did the one at Harvard. Accordingly, a greater variety of orientations was presented to the members and the atmosphere was, so to say, emotionally milder. The Harvard Workshop had a more unified and explicit dominant orientation, and a group process that was more heavily centered around the person of the responsible instructor; its members were actively responding to a personality as well as a point of view, and this emotional aspect led to increased learning in some cases, to non-learning in others. The diversity of orientation and of personnel in the Columbia setup produced less intense and less focussed affect, and in consequence a less threatening situation for the more authoritarian members. At the same time, however, such a group process is less likely to induce great involvement or ideological change in the Middles and Lows, if the present data have general validity.

The Columbia group showed insignificant changes on four of the six scales, but very great change on the Ethnocentrism Scale (see Table 3). These facts suggest that the Columbia Workshop was heavily focussed on specific intergroup problems, and perhaps that it was concerned primarily with information and skills and secondarily with personal-social orientation and insight. In the Harvard Workshop the emphases were just the reverse. I doubt that these two workshop functions can be given equal and strong emphasis in any single workshop. Which functions a workshop staff chooses to emphasize is, it would seem, a matter of personal predilection and training; which is the better arrangement in terms of over-all effects, we do not yet know.

## F. The Phenomenology of Change

Our study of workshop effects has thus far been centered on changes in scale scores, that is, on changes in a few of the many personal-attitudinal characteristics which the workshop staff attempted to modify. One of the many other forms of possible change which merit study is what we might call the "phenomenology of change." What meaning has the workshop had to its members? What ideas and principles loom largest in their thinking at the end of the workshop, as compared with the beginning? Only a limited step in this direction was attempted here; I wish to mention it briefly as a promising lead for future research. In the first group meeting, the members were given two or three days in which to write informal, several-page papers on the following three questions.

> 1. How have your ideas about minority problems developed? Starting with childhood, describe some of your major experiences with members of other religious or cultural groups. What were your early atti-

> tudes toward your own group and others, and how did these change, if at all?
>
> 2. What major intergroup problems are you now concerned with in your work or community, and how would you like to deal with them?
>
> 3. What would utopia be like, with regard to intergroup relations? Broadly speaking, what are the best ways of achieving this ideal state of affairs? What are the difficulties?

At the end of the workshop a similar procedure was followed. The only change was in the first question; the workshoppers were now asked to evaluate the workshop experience, to indicate how, if at all, their approach to intergroup problems had changed during the course of the workshop. Unfortunately, many of the participants had to leave Cambridge the day following the last session and were extremely busy during the final week. Only half of the group carried out the assignment, and we were therefore unable to perform a systematic analysis of the papers.

In the absence of formal data, a few general observations on the papers may be of value. Some of them were rather impersonal statements regarding the causes and cures of discrimination. The majority, however, were organized around issues and problems of specific personal relevance to the writer. These papers gave evidence of an active attempt to relate workshop material to one's personal, inner, and external problems. The writers were often unready to change their views on certain issues which had been the subject of some reading and discussion. It is, nevertheless, most rewarding to find students searching for personally valid viewpoints—and, equally important, being able to communicate these efforts to the instructor in writing—whether or not their views agree with his own. Here is an example.

> One workshop member wrote a good deal about his personal history. He is descended from a long line of New England Yankee Protestants and, although in his youth he rebelled against the conservative, traditionalistic views of his family, he is still consciously concerned with problems of status and class position. Prior to the workshop, he believed that Negroes were reasonably satisfied with their conditions of life; there were certain gross injustices to be remedied, but in the main Negroes wanted to go their own separate way without major change in their general societal position. Now, at the end of the workshop, he believes that discrimination is a more serious problem than he had thought; he feels that a fuller integration of Negroes into the mainstream of American life is desired by Negroes themselves, that this is a legitimate aspiration and one worthy of support. As a teacher in the area of the Humanities, he plans to assign readings that combine academic-literary value and contemporary social relevance, and he will encourage discussion of "human relations" problems. He ex-

> pects also to be more active within the school system generally, to reduce discrimination at both the student and faculty levels. Nevertheless, he is unprepared "to go all the way" and he discusses this honestly and at some length in relation to his concern with his own social position. For example, he feels that, in principle, the renting and owning of houses should operate on a completely nonsegregated basis. He would not join any attempt to exclude Negroes from living in his neighborhood. At the same time, he is not ready to live in a neighborhood containing many Negro families because, as he puts it, this would be too great a threat to his already insecure status. Nor is he ready to become a leader in "fighting for the downtrodden," for this might also estrange him from those higher-status groups in which he is even now a marginal member because of his occupation. This is not the document of an authoritarian personality, and certainly not of a prejudiced one. He recognizes the injustice of a viewpoint which he is nevertheless unable fully to oppose, and he is aware of a conflict without being able as yet to resolve it. It is difficult to predict what direction his adjustment will take in the long run. For the moment, however, I am inclined to count him among the workshop "successes."

Perhaps the most widespread phenomenological change is in the sense of militancy: the idea that there are urgent human problems to be met, that each individual must do what he can to change himself and others, and that action on many fronts—in the school, in the family, in industry, in the environments of all groups, and through legislative action—is required. The sincerity and conviction expressed in some documents offers additional assurance that the workshop had genuine psychological effects on its members.

A final word is in order concerning a crucial but neglected problem in this (and most other such) research. How durable are the changes wrought, and to what degree will they be translated into socially effective action? Our study has been limited to the inner dispositions, the ideas and feelings of our subjects. Yet, though it is true that "action begins in the minds of men," we must also bear in mind that "faith without works" is of at best limited value.[3] A follow-up study was omitted from the present research only because it was not feasible. In the absence of controlled data, it should be noted that some changes at the action level have occurred. These include school policy changes by our superintendent, curriculum changes by various teachers, and active membership in various educational and civic organizations by a number of workshop members.

---

[3] I am indebted to Dr. Herbert Conrad for reminding me, in a way I could not ignore, of the significance of this ancient wisdom.

## G. SUMMARY AND CONCLUSIONS

The aims of this study are: (*a*) to formulate in broad outline the *major functions* that Intergroup Relations workshops are intended to fulfill; (*b*) to determine the *effectiveness* of a particular workshop—one given at Harvard University in the Summer of 1951—in carrying out these functions; and (*c*) to discover some of the *determinants* of change in the group as a whole and in various subgroups within it.

The intended functions of the Harvard Workshop can be grouped under three main headings. These functions are probably common to most workshops in this field, although there are great variations in emphasis and procedure. First, there is the *knowledge-skill* function: to provide facts and ideas concerning various aspects of group life, and skills for dealing with everyday "intergroup" problems. Second, there is an attempt to induce *emotional-ideological change:* to help the members clarify the meaning of concepts such as "prejudice," "discrimination," and "democratic group functioning," and to further their growth toward democratic modes of thinking and relating to self and others. Finally, the workshop has the function of providing an *intense, supportive living experience* for each participant, by means of satisfying relationships with others of widely differing socio-religio-ethnic backgrounds. These functions are discussed in relation to workshops generally and the Harvard Workshop in particular.

The empirical study of workshop effectiveness deals primarily with the second of the above-mentioned functions. In order to estimate the extent to which the workshoppers became more democratic and flexible in their general life orientation, six scales were administered at the start, and again at the end, of the six-week workshop session. The Scales are designed to measure Authoritarianism (*F*), Ethnocentrism (*E*), Religious Conventionalism (*RC*), Politico-Economic Conservatism (*PEC*), Traditional Family Ideology (*TFI*), and Internationalism-Nationalism (*IN*). In earlier studies the *F, E, TFI,* and *IN* Scales intercorrelated to a sufficient degree (.6-.7) so that they may be regarded as measuring various facets of an inclusive "authoritarian" syndrome. The *RC* Scale correlates in the neighborhood of .5 with these scales; their relation to the *PEC* Scale is more variable, the average correlation being .3-.4. A high score on any scale presumably indicates strong support of the orientation denoted by the scale title, a low score the opposite. Scores on each scale can vary between 1.0 and 7.0 points, with 4.0 as a kind of neutral point.

In comparison with a heterogeneous group of College Students (Sophomores in a large urban university), the Harvard Workshop on its initial testing exhibits little prejudice (low *E* Scale mean and truncated range). Its means on the Authoritarianism, Traditional Family Ideology, and Nationalism Scales, though still on the "democratic" (low-scoring) side of the center, are similar to those of the College Students and indicate that some workshop members have considerably stronger authoritarian tendencies than their gross stand on minority problems might suggest. The workshop group is politically more liberal than the College Students but shows a higher degree of religious conventionalism. Thus, although a number of workshop members take a consistently democratic position in the various ideological areas studied, there are also a number whose views are fairly consistently autocratic. This latter group tends to reject ideas that reflect punitive or vindictive sentiments toward conventionally "deviant" groupings (Negroes, foreigners, criminals, unruly children). However, it is receptive to subtler forms of prejudice and stereotyped moral judgment, to a nationalistic conception of international relations, and to a conformist, submissive, impulse-rejecting personal orientation.

The scales were also given to a Columbia Workshop on Intergroup Relations and a Control Group (a graduate Education class at Harvard) at the beginning and again at the end of the summer session. The profile of scale means for the Columbia group is very similar to that for the Harvard Workshop, but the former earns slightly lower means on most scales and appears to have fewer authoritarian members. The means for the second Harvard (control) group are sufficiently similar to those of the Workshop to justify its use as a control group in the assessment of Workshop change.

The correlations among the six scales in the initial testing of the Harvard Workshop are of the expected magnitude. The *F, E, TFI,* and *IN* Scales intercorrelate at the level of .6-.7; they correlate with *RC* in the neighborhood of .6, and with *PEC,* .3-.4. These correlations provide further evidence for the existence of contrasting authoritarian and equalitarian syndromes, as well as a less well-defined "middle" group.

The correlations between initial and final scores of the Harvard Workshop are .69 for the *PEC* Scale and between .87 and .93 for the remaining scales. These correlations indicate that, with the exception of *PEC,* all scales have adequate test-retest reliability and our subjects have pretty much the same *rank-order* (though not necessarily the same *absolute*) positions on the final testing as they did initially.

With regard to the average amount of change shown during the six-week period, it was predicted that the Harvard Workshop would show a significant downward ("democratic") shift on all scales, and that the Control Group would not. For the Columbia Workshop, significant change was predicted on the Ethnocentrism Scale, but no predictions were made regarding the other scales. The predictions are in the main confirmed. The Harvard Workshop exhibits a significant drop on all scales except *PEC*. The Control Group shows a decrease on two scales, an increase on two, and no change on the remaining two. The Columbia Workshop undergoes a consistent downward shift on all scales, but the change is significant at the 5 per cent level or better only in the case of *E* and *TFI*.

For the analysis of differential change within the workshop, the Harvard sample was divided into three "authoritarianism" groupings: High Third, Middle Third, and Low Third of the initial *F* Scale distribution. The Middle Third shows the greatest over-all drop (.54) on the six scales, while the most authoritarian (High) group changes the least (.18). The change of .41 for the Low Third almost equals that for the Middles. Thus, in the Harvard Workshop at least, the members' receptivity to new ideas is strongly influenced by their initial personality characteristics. A similar analysis was carried out for the Columbia Workshop. This group differed from the one at Harvard in two main respects: it showed less over-all change on the six scales (.22 as compared to .38); and the High Third in initial authoritarianism showed the largest drop in scale scores (.35), while the Low Third changed least (.11). Thus, in both workshops scale change is a function of degree of authoritarianism, but the function differs from one workshop to the other. The Harvard workshop had considerable influence on the Middle and Low groupings, but only slight influence on the Highs. The Columbia workshop, on the other hand, induced moderate change in the High grouping, but only slight change in those initially Middle or Low in authoritarianism. A rationale for the two sets of findings is presented and the implications for workshop procedures are discussed.

In order to supplement the scale data and to achieve greater insight into the meaning of the workshop experience, we asked the members to write personal documents about the workshop and some of the problems discussed in it. Although these protocols were not subjected to systematic analysis, they did provide a number of valuable impressionistic observations. In particular, they indicated some of the general "feeling" changes (militancy, enthusiasm, self-awareness, and the like) and the forms of compromise between earlier views and workshop-derived principles, that the scale data in themselves cannot show.

If our findings and interpretations have any validity, workshop staffs must take greater account of the kinds of personalities in which they are trying to induce change. The workshop must take a multi-faceted approach, one that can capture the imagination and enthusiasm of its more receptive participants and at the same time not alienate those members who have difficulty in looking deeply or critically at themselves and their society.

A final word concerning the value of these workshops. To those who expect, or claim, that workshops as now constituted produce drastic changes in most participants, the findings of this study will perhaps be disappointing. There are, however, several grounds for taking a happier view. The mere existence of the workshop within the university tends to legitimize the democratic principles on which it is based. The workshop members are of strategic importance in their occupations and community organizations; any positive change that the workshop induces—in raising their morale, in broadening their understanding, in sensitizing their perceptions, in furthering their educational skills—is well worth the effort expended. Workshops appear to be doing this to a sufficient degree to justify their existence. They are no panacea, but they do help.

## REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. The Authoritarian Personality. New York: Harper, 1950.
2. ALLPORT, G. W. The Resolution of Intergroup Tensions. Pamphlet published by the Nat'l Conf. of Christians and Jews, New York, 1952.
3. COOK, L. A. (*Ed.*). College Programs in Intergroup Relations. Washington: American Council on Education, 1950.
4. DEUTSCH, M., & COLLINS, M. E. Interracial Housing. Minneapolis: Univ. Minnesota Press, 1951.
5. LEVINSON, D. J. An Internationalism-Nationalism Scale and some of its sociopsychological correlates. (Report in preparation.)
6. LEVINSON, D. J., & HOFFMAN, P. Studies in personality and ideology: Theory and measurement of Traditional Family Ideology. (Mimeographed paper, 1952.)
7. LEVINSON, D. J., & LICHTENBERG, P. Studies in personality and ideology: Theory and measurement of Religious Conventionalism. (Report in preparation.)
8. LEVINSON, D. J., & SCHERMERHORN, R. A. Emotional-attitudinal effects of an intergroup relations workshop on its members. *J. of Psychol.*, 1951, **31**, 243-256.
9. LIPPITT, R. Training in Community Relations. New York: Harper, 1949.
10. MCNEMAR, Q. Psychological Statistics. New York: Wiley, 1949.
11. ROSE, A. (*Ed.*). Race Prejudice and Discrimination. New York: Knopf, 1951.

*Department of Social Relations*
*Harvard University*
*64 Plympton Street*
*Cambridge 38, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 127-130.

# PERCEIVED NEGRO GROUP BELONGINGNESS AND SOCIAL REJECTION*

*Department of Psychology, Philander Smith College*

---

MARTIN M. GROSSACK

---

## A. THE PROBLEM

Belonging to the Negro group for a given individual may be accompanied by a variety of perceptions, expectations, motives, and defenses. A Negro may be "proud to belong" or ashamed, feel his position is a challenge or a handicap, prefer to "pass" into white society or to share his fate with that of other Negroes.

What are the socially shared norms, both ideal and real, in Negro life? It is assumed here that Negro group members tend to value characteristics of positive group belongingness as a normative ideal. Such an ideal individual has group pride, consciousness of his group's problems, and shares the cultural heroes and aspirations as well as the social disadvantages of other Negroes.

If such an ideal operates in Negro life, those who typify it will be socially approved and accepted, while deviates from the norm would experience rejection and disapproval. This rationale suggests three hypotheses:

1. Attraction to the Negro group (or belongingness) is an ideal social norm shared by Negroes.
2. Individuals typifying positive Negro group belongingness will be socially accepted by Negroes.
3. Individuals deviating from the norm of positive Negro group belongingness will be socially rejected by Negroes.

## B. PROCEDURE

To test these hypotheses, a hypothetical social situation was created in story form. The story had four variations in theme: (*a*) positive group belongingness, moves towards others; (*b*) positive group belongingness, moves away from others; (*c*) indifferent to group, moves towards others; (*d*) indifferent to group, moves away from others.

---

*Received in the Editorial Office on February 21, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

### 1. *Variation A*

This story is about Dick Chandler, resident of Cottonville, U. S. A. Dick is of average height, intelligent, and friendly. He is proud to be a Negro, active in *NAACP,* and opposes segregation. Dick reads "Ebony" and "Crisis" and admires Ralph Bunche.

The Supreme Court has ruled against segregation in public schools. However, the town of Cottonville has not taken steps to integrate Negro and white schools. Negroes in Cottonville became upset and formed a picket line, protesting in front of the City Hall. Dick joined the picket line.

### 2. *Variation D*

This variation, except for the following substitutions, was identical to *A*.

. . . intelligent and friendly. He feels pretty bad that he is a Negro, keeps away from the *NAACP,* and doesn't mind segregation. Dick doesn't read the Negro press and dislikes Ralph Bunche.

. . . in front of City Hall. Dick came by, decided he better have nothing to do with this and walked away.

The other variations were identical to the ones above, having the other two possible combinations of thematic content.

One hundred and eleven Negro student *S*s were randomly given one of the four hypothetical stories to read, then asked seven questions. These included the five described in Table 1 and the following two: (*a*) Was Dick right in what he did when he saw the picket line? (*b*) Would you have joined the picket line?

## C. RESULTS AND INTERPRETATION

Table 1 presents results to the post-stimulus questionnaire showing differences between recipients of positive group belongingness ("proud" Dick) and indifferent to the group ("indifferent" Dick) themes.

These findings substantially support the hypotheses of this research. The fictional character with positive group belongingness was socially accepted and approved very significantly more than the indifferent one. *S*s were more likely to want "proud" Dick as a fellow group member, roommate, friend, and to feel friendly toward him. Also, the individual portrayed as having positive group belongingness was more often perceived as having a healthy personality.

No differences were found to occur in response to the themes of joining or shying away from the picket line. Subjects were about equally divided on the moral righteousness of joining or shying away from the picket line and whether or not they would have joined it.

The results in Table 1 lead to the inference that positive group belongingness is a socially shared norm among Negroes. The experimental procedure seems to have powerful possibilities for the testing of social-psychological hypotheses. By presenting hypothetical social situations, one can get at a variety of variables such as conformity-deviance, acceptance-rejection, etc.

TABLE 1

QUESTIONNAIRE RESPONSES AS INFLUENCED BY STIMULUS STORY

| Item | Story received | Response<br>Much | <br>Little | $X^2$ | P |
|---|---|---|---|---|---|
| 1. To what extent would you like Dick as a friend? | "Proud" Dick | 43 | 12 | 36.21 | .001 |
| | "Indifferent" Dick | 15 | 41 | | |
| 2. To what extent would you like to have Dick as a roommate? | "Proud" Dick | 41 | 13 | 28.60 | .001 |
| | "Indifferent" Dick | 14 | 42 | | |
| 3. To what extent would you invite Dick to join a group you belonged to? | "Proud" Dick | 39 | 15 | 8.56 | .01 |
| | "Indifferent" Dick | 23 | 29 | | |
| 4. How friendly do you feel toward Dick? | "Proud" Dick | 41 | 14 | 8.18 | .01 |
| | "Indifferent" Dick | 27 | 29 | | |
| | | Yes | No | | |
| 5. Does Dick have a healthy personality? | "Proud" Dick | 46 | 9 | 31.24 | .001 |
| | "Indifferent" Dick | 17 | 38 | | |

Results of Festinger, Schachter, and Back (1) and of Schachter (3) are substantiated here. In their study of a housing community, the former writers found a close association between deviance of attitude and sociometric rejection. Schachter experimentally found that paid participants, deviating from the opinions of others in a group, were most often rejected sociometrically. Rejection of deviates was more frequent in groups of high cohesiveness and situations where the deviant behavior occurred in a group discussing material central to its goals. Similarly, in the present research, deviation from positive group belongingness is dysfunctional for the Negro group and social rejection could be expected.

A review of the literature on group belongingness and its significance for minority groups (2) suggests that the self-hatred, disillusionment, and indifference in such groups has been overstressed. This research supports that point of view.

## D. Conclusions

Hypothetical stories were developed presenting Negroes as having positive and indifferent group membership characteristics. A post-story questionnaire answered by 111 Negro *S*s supported hypotheses of belongingness as a Negro social norm and predicted rejection of deviates and acceptance of conformers to the assumed social standard.

## References

1. Festinger, L., Schachter, S., & Back, K. Social Pressures in Informal Groups. New York: Harper, 1950.
2. Grossack, M. Group belongingness and authoritarianism in Southern Negroes. (To be published.)
3. Schachter, S. Deviation, rejection, and communication. *J. Abn. & Soc. Psychol.*, 1951, **46**, 190-207.

*Department of Psychology*
*Philander Smith College*
*Little Rock, Arkansas*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 131-137.

# RELATIONS BETWEEN THE SELF-IMAGE, THE IMAGE OF THE FOREIGNER, AND THE DESIRE TO LIVE ABROAD*[1]

*Center for International Studies, Massachusetts Institute of Technology*

---

HOWARD V. PERLMUTTER

---

## A. INTRODUCTION

While the growing number of studies of national stereotypes provides interesting summaries of the traits different individuals agree are "typical" of Germans, English, Russians, etc., we know little of the *adaptive significance* of these characterizations for each individual. For if the referent of cognition is as unclear as the "typical German" or the "typical Italian," we might expect that some of the individual's projective mechanisms would exert their influence on the image formed. The image of the foreigner under some conditions mirrors the image of the self.

While the studies of Queener (3), Newcomb and Svehla (2) and others have contributed to the notion that family and college reference groups determine "international attitudes" (and by implication, images of foreigners), they cannot and do not profess to account for the idiosyncratic meaning a specific foreign image possesses.

The authoritarian studies of Adorno *et al.* (1) come closer to our interest because they present data concerning the Hi-F scale individual's predisposition to see foreigners, or more precisely, minority group members, as "disreputable and undesirable." These events are in part accounted for in terms of the Hi-Authoritarian's tendency to assign disliked characteristics to others, which he himself possesses but cannot accept. The image of the Jew and the Negro, then, serves some defensive system oriented toward achieving a psychologically stable world.

The particular individual's image of the typical foreigner depends as well on a variety of vicarious contacts with foreigners, in caricatures in movies, in radio and television, as well as on direct observation of foreigners here and abroad. This information, vast and variegated, must be assimilated by selection and learning, in accordance with principles not different from other

---

*Received in the Editorial Office on February 22, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This is one of a series of studies of "Personality Determinants of International Communication" being conducted in the International Communications Program at the Center for International Studies, Massachusetts Institute of Technology.

everyday experiences concerning individuals, direct contact with whom is limited. The individual, in improving or at least adding to his description of himself and others, can categorize the foreigner with familiar, personally meaningful concepts that must be in some respects similar to and different from himself, as he sees himself.

Since the purpose of this study was to determine what correlations exist between images of the actual and the ideal self, the relative desire to travel to a specific country and the image of the "typical member" of that country, it is appropriate that we consider briefly the psychological nature of these variables. In this fashion, some "predictions" concerning possible corrections can be made.

## B. THE ACTUAL AND IDEAL-SELF IMAGE

If we ask an individual to categorize the data he has available on himself, he must report on his distillation of the myriad of self observations, clear and hazy, consistent and inconsistent, reportable and private. Studies such as Sears (4) have shown that an individual can select from a series of trait names or activities those which characterize "himself."

Furthermore, most individuals can conjure up an image of what they would like to be, by mustering a series of traits they feel they *would like* to possess in greater or lesser degrees. Thus the individual can wish that he were more "mature," more "democratic," etc.

From one angle, at least, this ideal-self characterization requires that the individual indicate which characteristics, possessed in different amounts *by others,* he would like to possess. This formulation, however exact, is useful in making a derivation which follows later.

## C. THE IMAGE OF THE FOREIGNER

If, on the other hand, we ask the individual to depict a "typical Frenchman" or "Britisher," we are requiring him to reduce that multitude of impressions of temperamental, economic, political, and historical associations with the French in terms of his repertory of traits. Furthermore, he is confronted with the task of depicting not only "typicality," but a nationality grouping *different* and *distant* from his own self and country, as well.

## D. THE DESIRE TO TRAVEL TO A SPECIFIC COUNTRY

Expression of the desire to travel is in some respects a more complicated situation. We are in one respect asking with what kinds of people the individual would like to have contact. Our first-approximation postulation is that a desire to travel, although manifested in a multitude of conscious and rational motives, is in most cases an autistic wish to move to a distant place to interact with specific categories of people.

## E. HYPOTHESES

Put in this way, a series of hypotheses emerged:

1. The greater the relative desire to live in a foreign country, the more likely that the image of the "typical member" of that country will correspond to the individual's self image. This is based on the notion that, on the average, individuals tend to prefer contact with those who are similar in most respects.

2. The greater the relative desire to live in a foreign country, the more likely that the image of the "typical member" of that country will correspond to the individual's ideal-self image. This is based on the notion that people tend to prefer contact with those individuals they admire, e.g., are ideal representations of themselves.

## F. PROCEDURE

The subjects were summer session students at Boston University, and hence had varying occupations and income levels. All were told that they were filling out a questionnaire with "potential importance for understanding international relations," and were encouraged to express their sentiments honestly and sincerely. Since their names were not put on the paper, their responses were assured of complete secrecy. The reaction to the task seemed to be generally enthusiastic. The questioning after the completion of the questionnaires indicated a high level of interest and concern with their judgments.

A standard list of traits, used in several previous experiments because of the variety and clarity of the items, was presented. The students first checked those traits they felt were possessed in large and small amounts by typical members of two of the following four countries: France, Germany, Great Britain, and Italy. The order was varied so that one individual might fill out his conception of the typical Frenchman and then the typical German, while another would fill out the typical German first and the Frenchman second. The rest of the questionnaire required the students to use the same list to report traits actually possessed in large amounts and those in negligible amounts. The ideal-self page asked the students to use the same trait list to report traits they *would like* to possess in large amounts and those in negligible amounts, also.

Finally, the students filled out a series of rating scales: (*a*) His relative desire to spend all of one year in one foreign country with all expenses paid. The countries ranked were France, Germany, Italy and Great Britain. (*b*) The strength of his desire to live abroad, from: No desire to live abroad, to Would very much like to live abroad. (*c*) The maximum *length of time* he would like to remain in each of the four countries.

## G. RESULTS

### 1. *Relations between the Image of a Country and the Self-Image*

It will be recalled that each subject characterized "typical members" of two different nationalities, in addition to characterizing himself as he believes he is and would like to be. The data were handled in this way:

*a.* We determined the number of congruencies between traits he said he possessed in large amounts and the traits he ascribed in large amounts to each of the two typical foreigners.

*b.* The relative ranking of the countries (from which the two typical foreigners come) on the scale measuring ranked preference to spend one year abroad, all expenses paid. If the subject ranked France higher than Italy, France was termed the *preferred* country for this group of subjects.

*c.* The number of actual self-traits coinciding with the less-preferred country member's description was subtracted from the number of actual self-traits coinciding with the traits of the member of the more preferred country. So if the subject's picture of a typical Frenchman had eight traits he himself possessed in large amounts, and his picture of the typical Italian had five traits he himself possessed in large amounts, the trait difference score would be +3.

*d.* The sign of trait-difference score was compared with the relative ranked desire to travel score. If the subject preferred Country *A* to Country *B*, and the sign of trait difference score of *A*-*B* was positive, the column labeled Plus was checked. If the sign of trait difference score was negative for the preferred country, the column labeled Minus was checked.

*e.* The same calculations were made for the ideal self congruencies with typical foreign members.

Table 1 summarizes the number of times the Plus Direction and Minus Direction columns were checked. The results give some support to both hypotheses. Hypothesis 1, however, receives consistent support, while a sharp reversal occurs in the French-German group. In this group subjects tended to assign more ideal self-traits to Germans, even though they preferred to go to France.

TABLE 1
Relation between Desire to Travel to a Country and Self-Country Image Congruencies

| Typical country members described | Number of subjects | Actual self | | Ideal self | |
|---|---|---|---|---|---|
| | | Plus direction | Minus direction | Plus direction | Minus direction |
| English and Italian | 25 | 15 | 7 | 12 | 11 |
| Italian and German | 24 | 15 | 7 | 11 | 9 |
| English and French | 29 | 16 | 8 | 21 | 5 |
| English and German | 16 | 12 | 2 | 11 | 3 |
| French and German | 25 | 12 | 9 | 6 | 17 |
| Totals | 119 | 70 | 33 | 61 | 45 |

## H. The Desire to Travel and Self Esteem

What relations could we expect between an individual's reported picture of himself and his desire to travel?

First, it was possible to distinguish two groups of individuals: (*a*) High self esteem individuals: those who tend to characterize themselves as having homogeneously positive traits, with *no* undesirable traits. (*b*) Low self esteem individuals: those who tend to characterize themselves as having a relatively large number of negative or undesirable traits.

Then it was possible to group our subjects along these dimensions: (*a*) Strong desire to travel individuals, that is, those individuals who give a high rating on the "want to go abroad" scale. (*b*) Low desire to travel individuals, that is, those individuals who give a low rating on the "want to go abroad" scale.

Table 2 indicates the mean number of negative traits possessed by individuals who have a strong, moderate, or no desire to travel. The results show a successive trend in the direction: the stronger the desire to travel, the greater the number of negative self-traits reported.

These traits were not only consistently ranked as traits the students would

TABLE 2
Relation between Relative Desire to Travel and Reported Negative Traits

| | Range of rank | *N* | Average number negative traits |
|---|---|---|---|
| Strong desire to travel | 4-5 | 29 | 2.93 |
| Moderate desire to travel | 3-4 | 44 | 2.19 |
| Little desire to travel | 2-3 | 22 | 2.18 |
| No desire to travel | 1-2 | 25 | 1.60 |

like to possess in negligible to zero amounts, but were also ranked by other groups as undesirable traits to possess. Examples are "immature," "hostile," "suspicious," "worrying," "dependent," "narrow interests," "depressed," "unimaginative."

Conversely, the average rankings of desire to travel for individuals who gave themselves *no* negative traits (high self esteem individuals) were compared with individuals who gave themselves *five or more* negative traits. Table 3 gives the results of this comparison.

TABLE 3
CHARACTERISTICS OF HIGH AND LOW SELF-ESTEEM INDIVIDUALS

| *N* | Category | Mean rank desire to travel | P | Mean total months abroad | P |
|---|---|---|---|---|---|
| 34 | High Self Esteem (zero neg. traits) | 2.6 | .05 | 37.0 | n.d. |
| 17 | Low Self Esteem (>5 Traits) | 3.4 | | 45.0 | |

The results of a test between the means of the two distributions indicate a difference at the 5 per cent level of confidence. While there is a trend in the predicted direction, low self esteem individuals, when asked how many months they would spend abroad in various countries, did not choose to stay significantly longer than high self esteem individuals.

## I. DISCUSSION

Since the results relevant to the hypotheses are in the predicted direction, the theory is seen to be supported. But the evidence is not sufficiently unequivocal to eliminate questions about other mechanisms that may be operating. The desire to travel abroad admittedly stems from a multitude of motives. It may be as much a need to cease interaction with contemporaries as a need to interact with foreigners. A positive image of the foreign member will then probably be less important than the existence of a negative image of in-group members. This notion seems in line with the results of the breakdown of the subjects into high and low self esteem individuals. The greater desire to emigrate by low self esteem individuals suggests that the need to emigrate may be a magical expression of a need to leave contemporary, in-group people and objects, e.g., self-connected things. However, there are other interpretations of the significance of the "desire to travel" motive. It is possible to conclude that the high self esteem individuals, who do not admit of *any* negative traits, must be defensively oriented, concerned with maintaining a homogeneously desirable picture of the self, while the low self esteem people are willing to accept and report some personal deficiencies. Future research is needed to determine more precisely the significance of the need to emigrate, its relation to the conception of the in-group, the self, and the ideal self.

## J. SUMMARY

One hundred nineteen students filled out descriptions of typical members of two of the following foreign countries: France, Great Britain, Germany, and Italy, then described themselves, what they would like to be, rated along different dimensions their desire to spend time in a particular country, and several other dimensions tapping interest in traveling abroad.

These two hypotheses posed were upheld:

1. The greater the relative desire to live in a foreign country, the more likely will the image of the "typical member" of that country correspond to the "self image."
2. The greater the relative desire to live in a foreign country, the more likely will the image of the "typical member" of that country correspond to the individual's ideal-self image. This relationship was not as strong as the previous.

A breakdown of students was made into those individuals with high self-esteem (reported they had *no* undesirable traits) and those with low esteem (reported five or more undesirable traits). The results show that the low self-esteem students show a significantly greater desire to travel abroad than the high self-esteem students. Interpretations of these findings were made.

## REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. The Authoritarian Personality. New York: Harper, 1949.
2. NEWCOMB, T. M., & SVEHLA, G. Intra-family relationships in attitudes. Sociometry, 1938, **1**, 180-205.
3. QUEENER, E. L. The development of internationalist attitudes directed toward peace. Grad. School, Yale University Ph.D. dissertation, 1947.
4. SEARS, R. R. Experimental studies of projection: I. Attribution of traits. *J. Soc. Psychol.,* 1936, **7**, 151-163.

*Center for International Studies*
*Massachusetts Institute of Technology*
*50 Memorial Drive*
*Cambridge 39, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 139-149.

# GROUP STRUCTURE AND THE BEHAVIOR OF INDIVIDUALS IN SMALL GROUPS*

*Department of Psychology, The Johns Hopkins University*

Marvin E. Shaw[1]

## A. Purpose

Several recent experimental studies of small group behavior (2, 3, 5) have utilized the Bavelas model (1) to specify the characteristics of the group structure and have attempted to demonstrate the relationship of these characteristics to the processes of group interaction. The results of these studies show that the structural measure which Bavelas called *centrality* is to some extent correlated with certain measures of group behavior. While these findings are sufficiently interesting to indicate that the Bavelas approach provides a useful framework for the study of small groups, the relationships between structural and behavioral measures are not as precise as would seem to be desirable.

The purpose of this paper is to re-examine the centrality measure in order to show some of its inadequacies as revealed by experimental evidence, and to present an alternate structural measure to be used for predicting the behavior of individuals in small groups.

## B. Centrality and Its Relationship to Behavior

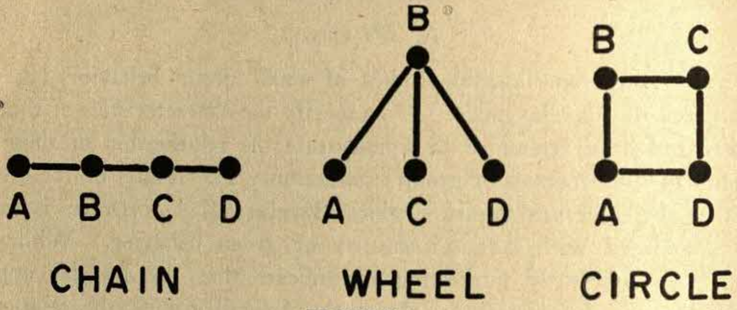
In the Bavelas system, group structures consist mainly of positions (or cells) connected by various arrangements of communication channels (or links). Some possible arrangements are shown in Figure 1. To compute centrality the first step is to compute a measure called *sum of distances.* This measure is obtained by counting, for each position, the minimum number of channels which must be crossed in order for that position to communicate with all other positions in the group, and then summing this value for all positions in the group. An example of this procedure is given in Figure 2. *Individual centrality* is then determined by finding the ratio of the sum of distances for the whole structure to the sum of distances for the particular position; *net centrality* is the sum of the individual centralities.

---

*Received in the Editorial Office on February 25, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was done under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is report No. 166-I-180, Project Designation No. NR-145-089, under that contract.

Bavelas (2) reported some preliminary attempts to apply his model to the study of problem solution by small groups. He concluded that there are differences among nets and that these differences are revealed most clearly with respect to (*a*) the location in the net of recognized leadership, (*b*) the probability of errors in performance, and (*c*) the general satisfaction of group members.



CHAIN WHEEL CIRCLE

FIGURE 1
Some Possible Arrangements of Communication Channels Among Four Individuals

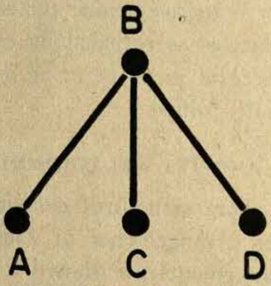


FIGURE 2
Calculation of Distances in the Wheel Pattern: The Sum of Distances from *A* to All Others is 5; from *B* to All Others is 3; from *C* to All Others is 5; and from *D* to All Others is 5; Total Sum of Distances is 18

In a more comprehensive study, Leavitt (3) investigated group problem solving in four, five-position nets. He concluded that centrality could be used to differentiate among positions in the net as well as among nets. More recently, Shaw (5) investigated three, four-position nets in which distribution of information was introduced as an additional variable. For purposes of this discussion, we shall be interested only in the control groups of that experiment in which information was distributed equally among the four members of the group. In the control groups he found that, centrality differentiated among positions within nets but not among nets.

More specifically, both Leavitt and Shaw found statistically significant differences among positions within nets with respect to: (*a*) the number of messages sent; (*b*) morale, or general satisfaction of group members; and (*c*) the probability of being selected (or recognized) as the leader. Statistically significant differences among nets were found with respect to: (*a*) the total number of messages; (*b*) morale, or total general satisfaction of group members; and (*c*) the emergence of leadership (defined as number of unanimous choices).

Leavitt concluded that all these measures of behavior were related to centrality, the possible exception being recognition of leadership which he believed to be a function of the maximum difference between individual centralities within the net. Shaw found that the centrality could be used to predict the behavior of individual positions but that this measure was unsatisfactory when used to predict for the net as a whole. Also, the emergence of leadership in the net seemed to be influenced by the number of positions having the highest individual centrality in the net as well as by the maximum difference between individual centralities. Both Leavitt and Shaw expressed the view that centrality is effective because of the degree of independence which it reflects.[2]

In Figure 3 the mean number of messages sent and the mean of the measures of morale have been plotted against individual centrality for the data reported by Leavitt and the data reported by Shaw. If individual centrality and the behavioral measures were precisely related, this plot should yield a reasonably smooth curve with no major inversions. The many inversions shown in Figure 3 suggest that individual centrality is an unsatisfactory structural measure for predicting group performance.

The same general picture emerges when other behavioral measures are considered. From the experimental evidence we can conclude that centrality: (*a*) allows qualitative statements about differences among individual positions in communication nets, but does not permit quantitative statements about these differences; (*b*) predicts differences among individual positions in the same net, but does not predict differences among individual positions in different nets (i.e., two positions having the same centrality index do not necessarily produce the same behavior, nor do two positions having different centralities produce different behavior); and (*c*) does not adequately reflect differences among communication nets.

---

[2] Independence is used rather broadly to refer to the degree of freedom with which a member of the group may function.

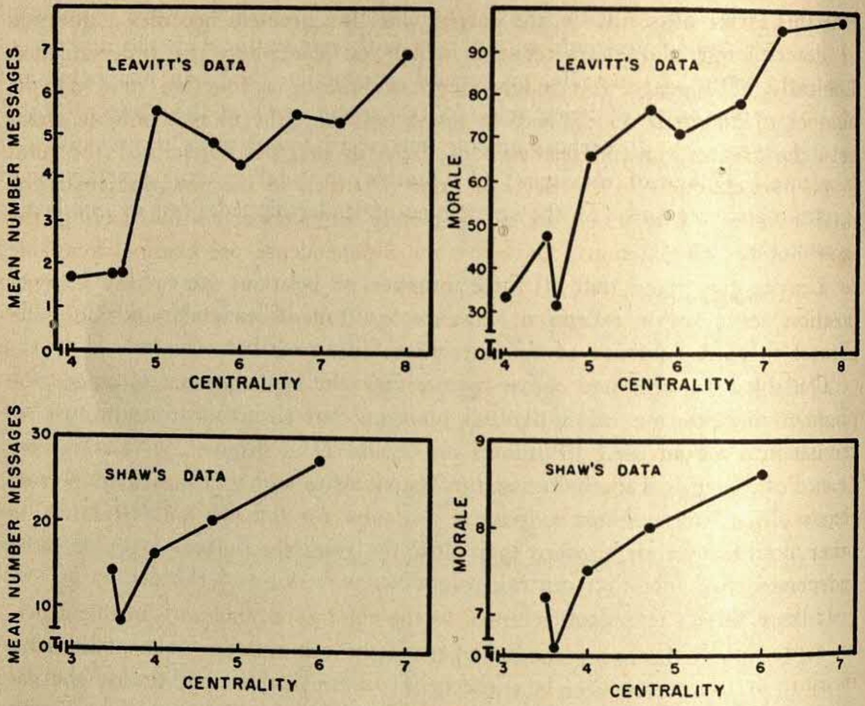


FIGURE 3
THE RELATIONSHIP OF CENTRALITY TO BEHAVIORAL MEASURES

### C. A SUGGESTED MEASURE OF INDEPENDENCE

Since experimenters who have attempted to correlate centrality with behavioral measures have agreed in interpreting the effectiveness of centrality as being due to the degree of independence associated with it, it seems worthwhile to consider how this is brought about. Conceivably, centrality could be directly correlated with the independence of the *S* assigned to the position. If this is the case, then the experimental results suggest that something other than the independence of the *S* contributes to the behavioral measures. This is very probably true, but such other factors, in a well controlled experimental situation, should be randomly distributed among the experimental conditions, thus contributing to variability, but not to mean differences.

Another possibility is that centrality is only partially correlated with independence. This would mean that there are some characteristics of the net and of positions within the net which help to determine the independence of the *S*s and which is not reflected in the centrality measure. Assuming that this latter possibility is the correct one, the problem becomes a question of determining the characteristics which do contribute to independence. Logically, it appears that independence should be a function of: (*a*) the number of channels available to a given position (the more available channels, the greater the independence of the *S* in that position); (*b*) the total number of channels in the net (the more channels in the net relative to the total number of channels in a completely interconnected net having the same number of positions, the lesser the independence per channel available to a given position); and (*c*) the number of positions for which a given position serves as a relayer of information (the more such positions, the greater the independence of the *S* in the relayer position).

The third factor listed above requires further explanation. Consider the plight of the position which depends upon another position for sending or receiving information (e.g., Positions *A, C,* and *D* in Figure 2). The relayer (e.g., Position *B* in Figure 2) has it in his power to withhold information from that position, or to refuse to pass on requests for information. Relative to other positions in the group, this certainly gives the relayer position more independence.

If these three factors are accepted as the chief determiners of independence, an expression is needed which will permit the computation of a quantitative measure of independence. It is the belief of the author that the number of channels available to a position will be the chief determiner of independence. This could be quantified by simply counting the number of channels available to each position in the net. The influence of the number of channels in the net can be taken into account by finding the reciprocal of the ratio of the total number of channels in the net to the number of channels in a totally interconnected net, and then multiplying this value by the number of channels assigned to the position under consideration.

This leaves the effects of serving as a relayer to be considered. Here we must distinguish between the situation in which the relayer is linked directly to the position for which he relays, and the situation in which the relayer is separated from the position for which he relays by one or more other positions. Because this particular factor is expected to contribute relatively little to independence, the effects of relaying for a position separated from the relayer by more than one position can be safely ignored. For the same reason, this factor cannot be quantified by simply counting the number of positions for which a position serves as a relayer. Empirically, the most satisfactory way of doing this is to convert the value obtained by the counting procedure to logarithms.

These considerations can be expressed more clearly by the following formula:

$$I = n + \left[ n \left( 1 - \frac{\Sigma n}{N} \right) \right] + \log R_d + \log R_i$$

where,

I = Independence

n = number of channels available to a given position

N = total number of channels in a completely interconnected net having the same number of positions

$R_d$ = number of positions for which a given position must serve as a first relayer

$R_i$ = number of positions for which a given position must serve as a second relayer.

## D. Relationship of Independence to Behavioral Measures

It is the purpose of this section to show how well the independence measure correlates with behavioral measures by making use of available experimental data. The general procedure was to compute the mean score for each posi-

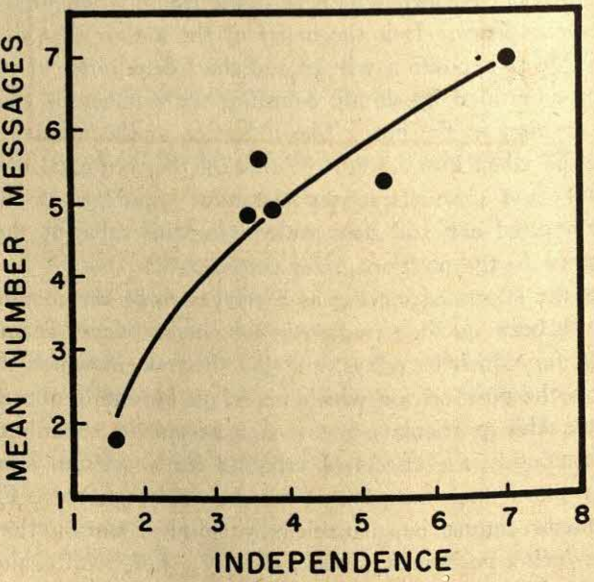


FIGURE 4
Mean Number of Messages Sent Per Problem as a Function of Independence (Leavitt's Data)

tion, and the grand mean for all positions having the same *I* value.[3] The mean score for each position was then plotted against the Independence measure, and the best-fitting straight line or curve was fitted to these plotted points, using the method of least squares. The *F* test described by Lewis (4) was used to determine goodness of fit. The function accepted was, in each case, the simplest one that could not be rejected as an adequate representation of the data.

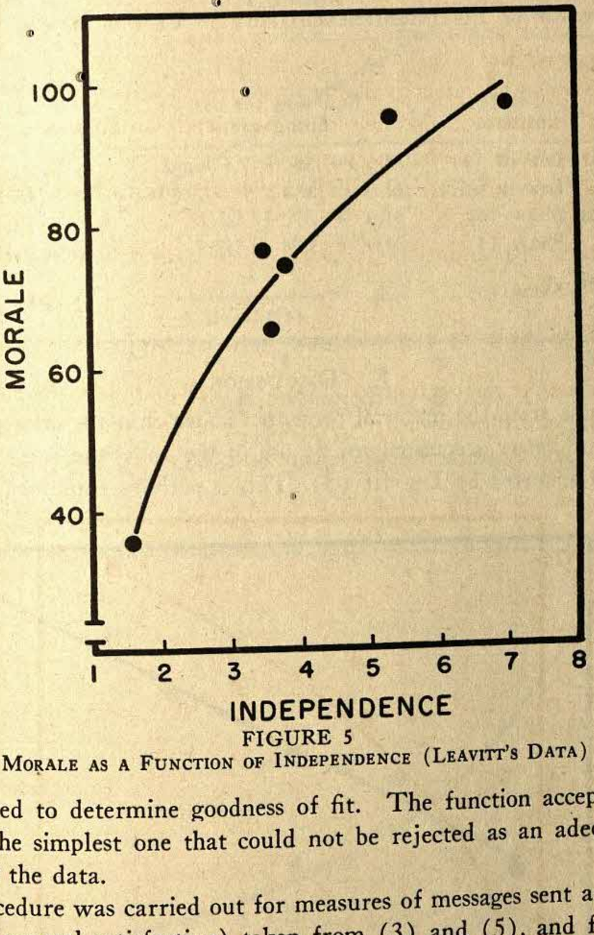


FIGURE 5

MORALE AS A FUNCTION OF INDEPENDENCE (LEAVITT'S DATA)

This procedure was carried out for measures of messages sent and measures of morale (general satisfaction) taken from (3) and (5), and for measures of recognition of leadership taken from (5). Figures 4, 5, 6, 7, and 8 were obtained by plotting the values computed from the fitted equations; the filled-in circles are the grand means for all positions having the same $I$ value. Table 1 gives the important statistics derived from the curve-fitting procedure.

---

[3] The measure of independence under discussion might be termed *individual independence* since it refers to individual positions within nets. One might suppose that *net independence* could be computed by summing all the individual *I* values in the net. However, this procedure would involve the assumption that individual independence and behavioral measures are linearly related. Since it is doubtful that this assumption is tenable, some other method must be devised for computing net independence. This problem is beyond the scope of the present paper.

TABLE 1
THE RELATIONSHIP OF THE INDEPENDENCE INDEX TO CERTAIN BEHAVIORAL MEASURES

| Behavior measured | Source of data | Equation for best-fitting curve | Coefficient or index of determination | Standard error of the estimate |
|---|---|---|---|---|
| Messages sent | Leavitt (3) | $Mes = .56 + 7.5 \log I$ | .62 | .52 |
| Morale | Leavitt (3) | $Mor = 18.12 + 97.02 \log I$ | .74 | 4.81 |
| Messages sent | Shaw (5) | $Mes = .97 + 5.2\, I$ | .95 | 1.00 |
| Morale | Shaw (5) | $Mor = 5.76 + .57\, I$ | .97 | .26 |
| Recognition of leadership | Shaw (5) | $\text{R.L.} = \dfrac{1}{.132 - .02\, I}$ | .95 | .87 |

## E. DISCUSSION

It is obvious from the material presented above that the data presented by Shaw (5) are better accounted for by use of the Independence measure than are the data reported by Leavitt (3). This is perhaps explained, in part, by

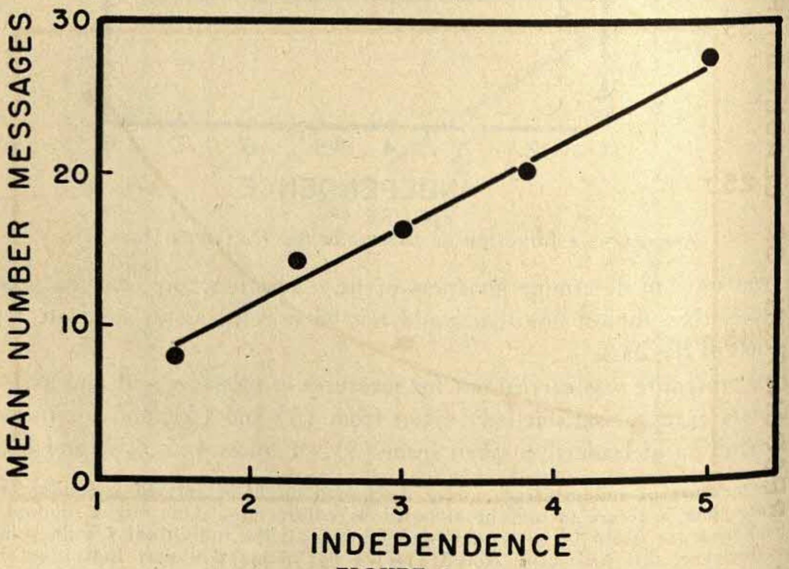


FIGURE 6
MEAN NUMBER OF MESSAGES SENT PER PROBLEM AS A FUNCTION OF INDEPENDENCE (SHAW'S DATA)

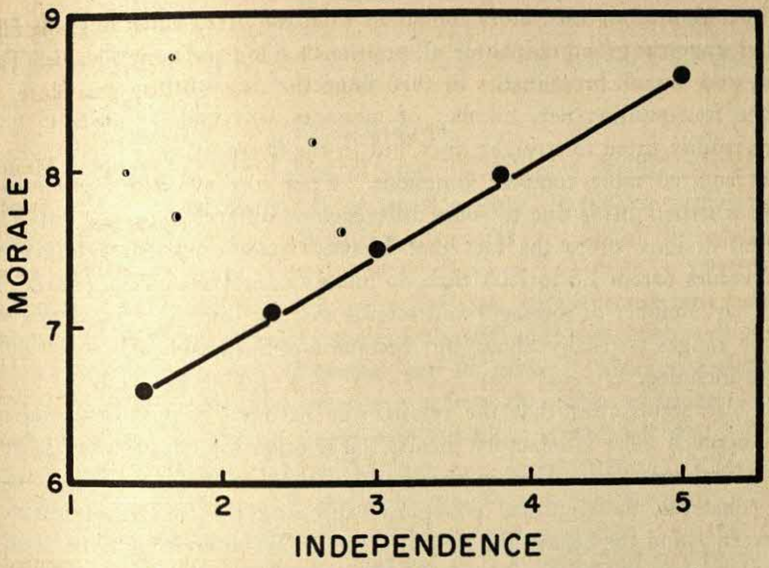


FIGURE 7
MORALE AS A FUNCTION OF INDEPENDENCE (SHAW'S DATA)

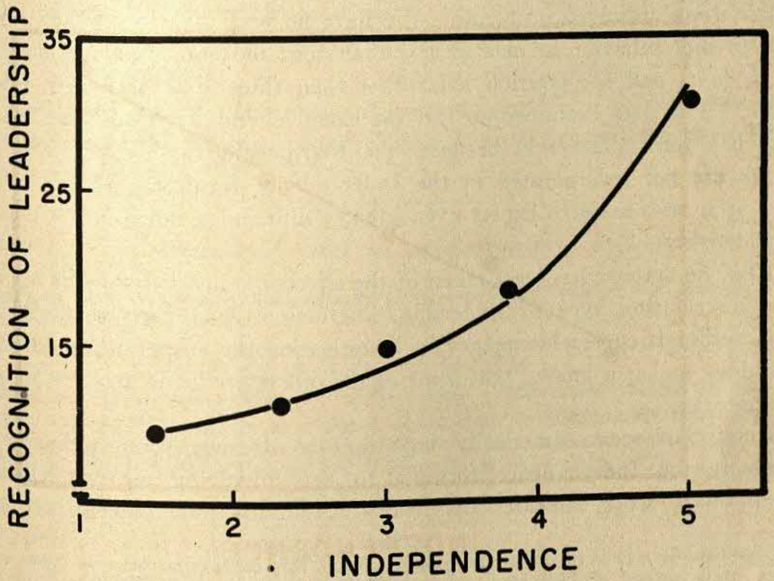


FIGURE 8
RECOGNITION OF LEADERSHIP AS A FUNCTION OF INDEPENDENCE (SHAW'S DATA)

the fact that Shaw used more replications (fifteen) than did Leavitt (five).[4] Other important factors include differences in *S* population, kind of problems used, and overall procedure. In this connection it is interesting to note that in the four-position nets number of messages sent and measures of morale were readily fitted to straight lines, but in the five-position nets similar measures required more complex functions. From the available data it is not clear whether this is due to some difference or differences in the two experimental designs, or to the fact that Leavitt's data represent a larger range of *I* values (from 1.6 to 7.0) than do Shaw's data (from 1.5 to 5.0). That is to say, number of messages and morale may be linearly related to *I* within the ranges tested by Shaw, but become non-linear when higher values of *I* are included.

It also seems clear that the relationship between *I* and behavioral measures is much more satisfactory than is the relationship between centrality and behavioral measures. Within the limits of the *I* values used in determining the constants, the empirical equations expressing the functional relationship between *I* and the behavioral measures should prove useful for the prediction of behavior in similar experimental situations.

It is believed that the development of the Independence measure represents some progress in dealing with this complex problem. There are, of course, some unsolved problems. We have no way of knowing how well *I* will predict behavior in nets of less than four or more than five positions, or in four- and five-position nets other than those that have been investigated. The data examined so far, for example, indicate that there are (or may be) some differences between the four-position and five-position nets which are not incorporated in the Independence formula. If this is true, then it is reasonable to expect even greater differences when larger nets are considered.

Also, no account has been taken of the effects of such variables as information distribution, type of leadership, instructions, and problem complexity. The author is currently engaged in investigating the effects of some of these variables, and it is known that some of the others are being attacked in other laboratories.

Finally, it seems practically important to discover some technique for applying the Independence formula to nets involving one-way channels. Preliminary work indicates that a distinction must be made between incoming and out-going channels when computing $I$. If later work proves this to be feasible, then the generality of application of the Independence measure will be greatly increased.

---

[4] The sample size is of particular importance in this connection because mean scores were used instead of individual scores. Variation among individuals treated alike was generally great; therefore, increased sample size should have a marked effect upon the reliability of the mean scores.

## F. SUMMARY

This paper examines the results of attempts to apply the Bavelas model for group structures, to the study of small groups. Particular attention is given to individual centrality, which is a structural measure of the ease with which an individual can communicate with all other individuals in the group.

It is shown that individual centrality does not permit quantitative statements about the differences among individuals within groups (although qualitative statements about such differences can be made). Since investigators agree that centrality is effective because it reflects the degree of independence with which an individual in a group may function, it was felt that a more direct measure of independence would overcome this deficiency. A formula for computing such a measure is presented which takes into account (*a*) the number of communication channels available to the individual, (*b*) the total number of communication channels available to the group, and (*c*) the number of individuals for which the given individual serves as a relayer of information.

This measure of independence accounts reasonably well for experimental measures of (*a*) morale (or general satisfaction), (*b*) number of messages used in solving problems, and (*c*) recognition of leadership.

It is believed that the Independence measure provides a sounder basis for the prediction of the behavior of individuals in small groups than those heretofore available.

## REFERENCES

1. BAVELAS, A. A mathematical model for group structures. *Appl. Anthrop.*, 1948, **7**, 16-30.
2. BAVELAS, A. Communication patterns in task-oriented groups. *J. Acoust. Soc. Amer.*, 1950, **22**, 725-730.
3. LEAVITT, H. J. Some effects of certain communication patterns on group performance. *J. Abn. & Soc. Psychol.*, 1951, **46**, 38-50.
4. LEWIS, D. Quantitative Methods in Psychology. Ann Arbor: Edwards, 1949.
5. SHAW, M. E. The effects of unequal distribution of information upon leadership, morale, and problem solution efficiency in various communication nets. Unpublished doctor's dissertation, University of Wisconsin, 1953.

*Psychological Laboratory*
*The Johns Hopkins University*
*1315 Saint Paul Street*
*Baltimore 2, Maryland*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 151-160.

# MODERN LOGIC AND TASKS FOR EXPERIMENTS ON PROBLEM SOLVING BEHAVIOR*

*Naval Research Laboratory, Washington*

Omar Khayyam Moore and Scarvia B. Anderson[1]

## A. Introduction

The resources of mathematical logic have been applied to a variety of problems in the empirical sciences. These applications frequently have taken the form of attempts to axiomatize the theoretical structures of some science or sub-section thereof, to explicate concepts and definitions, or to employ logical formulae as rules for testing the soundness of deductions made in a "natural" or "folk" language (e.g., English).[2]

Generally speaking, the techniques and findings of modern logic have been of most direct concern and utility to theoreticians. In this paper we shall describe a means by which mathematical logic may be of immediate use in experimentation. Specifically, we shall show how one of the calculi of modern mathematical logic, the calculus of propositions, may lend itself as a source of tasks for experiments on human problem solving behavior.

Before giving a detailed explanation of our approach and in order more adequately to differentiate it from other uses of mathematical logic, we shall indicate something of the nature and range of previous applications by reference to specific studies.

McKinsey, Sugar, and Suppes (13) have recently reformulated the axiomatic foundations of classical particle mechanics. Following the lead of Bloomfield (3), Bloch (2) has constructed a set of postulates for phonemic analysis within the field of linguistics. Greenberg (6), an anthropologist and linguist, has begun an axiomatic treatment of kinship systems. In psy-

---

*Accepted for publication by Ernest R. Hilgard of the Editorial Board, and received in the Editorial Office on February 27, 1954. Copyright by The Journal Press.

[1] The authors are indebted to Sherman Ross, Frank V. Taylor, and William D. Garvey for helpful suggestions from the psychological point of view and to Richard S. Rudner for valuable criticisms with respect to logic.

This article is based on work done under contract with the Office of Naval Research, Navy Contract Number N-onr-49403. The work was carried out at the Naval Research Laboratory, Washington 25, D. C., in what is now the Systems Division. The opinions and assertions contained herein are the private ones of the writers and are not to be construed as official or reflecting the views of the Navy Department or the naval service as a whole.

[2] This, of course, does not constitute a mutually exclusive or exhaustive classification.

chology there is, of course, the axiomatization of rote learning theory by Hull *et al.* (8). The examples given, and there are others which might have been mentioned with equal propriety (12, 22, 25), are characterized by relatively complete employment of the axiomatic technique. In each case an attempt is made to create a formal deductive system capable of a fruitful empirical interpretation within the domain of phenomena under consideration.[3] Of course, it is not always necessary to employ the full axiomatic technique. Carl Hempel (7), among others, makes excellent use of the techniques of logic to explicate methodological concepts in the empirical sciences. Moore (14) and Moore and Lewis (15) have worked primarily with the substantive concepts (e.g., "culture") of particular sciences.

Finally, a great many scientific workers make use of various logical formulae, such as Modus Ponens and the DeMorgan Laws, to check the validity of deductions which need not form any part of a systematic body of theory.

## B. PROBLEM

In the field of learning, there have been comparatively few experimental studies of what has been variously called "thinking," "creative thinking," "reasoning," and "problem solving."[4] In order adequately to study higher order problem solving behavior there is a pressing need not only for a body of testable theory, but also for a series of standardized tasks of varying levels of complexity.[5]

Our problem was (*a*) to determine whether there was at hand a source of tasks rich enough to yield a multitude of problems of varying levels of complexity, solvable through the use of one basic set of instructions and readily adaptable for many experimental purposes; and (*b*) if such a source of tasks was available, to effect an adaptation for use in studying human problem solving behavior.

## C. CALCULUS OF PROPOSITIONS

The calculus of propositions, also called the "Sentential Calculus," with which we are concerned as a source of tasks, is one interpretation of abstract

---

[3] Ideally, such a system would be comprised of a set of consistent, complete, and independent postulates together with theorems derived therefrom in strict accordance with determinate rules of procedure. All materials entering into the construction of such a system would be explicitly listed. There are wide differences in the degree to which various ones of the examples mentioned approach this ideal.

[4] Excellent discussions of problems involved in the study of phenomena of this order are provided by Humphrey (9), Johnson (10), Vinacke (21), and Wertheimer (23).

[5] Underwood (20) explicitly develops this point.

Boolean algebra. Excellent accounts of it can be found in a number of symbolic logic texts (1, 4, 18).[6]

The propositional calculus is a language, ideographic in notation, consisting of:

1. Three kinds of symbols: propositional variables (e.g., *“p,” “q,” “r”*), constants (e.g., “∼” “v,” “¬,” “/”),[7] and punctuation (e.g., “(,” “)”).
2. Formation and transformation rules governing respectively what combinations of symbols are permissible and what constitutes “proof” in the language.
3. A set of postulates and derived theorems.
4. Definitions (which, as in all deductive systems, may be eliminated).

There are many ways in which the calculus might be utilized as a source of experimental tasks. One option, for example, might involve (*a*) teaching subjects the formation and transformation rules, (*b*) presenting the postulates of the system, and (*c*) giving the subjects the task of deducing various theorems. Subjects would be engaged in the deduction of tautologies from tautologies. However, the experimenter would have more latitude in devising problems if the method of “conditional” or “hypothetical” proof were employed, for it would then be permissible to introduce hypotheses into the “initially given” portion of the problems. It was decided for this and other reasons to use one variation of “conditional proof” or “derivation” in our adaptation of the calculus. In order to facilitate the description of the work actually done and the advantages the calculus offers, an example problem, one used with subjects in our laboratory, is presented. In Sector I of the problem shown in Table 1, the subject finds the premise or premises with which he is to work. There are four premises in this particular problem. In Sector II is the answer he seeks. His problem is to determine whether, through the use of rules he has learned, he can deduce the expression given in Sector II. The information in Sectors I and II, as well as the rules, is provided by the experimenter. In Sector III, the subject records the steps which he takes in attempting to deduce the expression given in Sector II. In Sector IV is recorded the “justification” for each step taken.

Using the example problem as a point of departure, we shall now con-

---

[6] There are other calculi of propositions; e.g., Fitch (5) has worked out a modal calculus which differs considerably from the Whitehead-Russell system with which we are concerned.

[7] In some formulations of the calculus the “∼” and “v” are taken as primitive and other constants are introduced by definition. The choice of which constant or constants to take as primitive and whether to introduce others by definition, like choices concerning certain other features of the system, depends upon the purposes for which the calculus is to be used.

TABLE 1
EXAMPLE PROBLEM (FORMAL)

| SECTOR I | SECTOR II |
|---|---|
| 1. $(p \vee q) \supset (r \vee s)$ | |
| 2. $(q \vee r) \supset t$ | |
| 3. $\sim t$ | $s$ |
| 4. $p$ | |

| SECTOR III | SECTOR IV |
|---|---|
| 5. $p \vee q$ | from 4, rule of Addition |
| 6. $r \vee s$ | from 1 and 5, Modus Ponens |
| 7. $\sim t \supset \sim (q \vee r)$ | from 2, Contraposition |
| 8. $\sim (q \vee r)$ | from 7 and 3, Modus Ponens |
| 9. $\sim q \cdot \sim r$ | from 8, DeMorgan Laws |
| 10. $\sim r$ | from 9, Detachment |
| 11. $\sim r \supset s$ | from 6, Definition of Conditional |
| 12. $s$ | from 11 and 10, Modus Ponens |

sider certain of the characteristics of calculus tasks insofar as they are relevant to individual problem solving experiments.

### 1. *Mode of Symbolization*

The symbols used in the example problem are those frequently employed by logicians. There is no reason, however, why an experimenter might not use other symbols (e.g., Chinese characters) so long as the logical form of the deduction is preserved. What might be called "'mode of symbolization," then, may be treated as an experimental variable. An investigation of the effects on performance of subjects of diverse ways of symbolizing the same problems would be of both theoretical and immediate practical interest.

### 2. *Content*

The letters, *"p," "q," "r," "s,"* and *"t,"* that occur in the example problem are propositional variables.[8] These variables may be thought of as place holders into which propositions may be inserted. This is a fact of importance to the experimenter, since he is free to create an indefinitely large number of problems which, although differing in content, have the same logical form. Thus, "content" may be used as an experimental variable. The problem shown in Table 2 possesses the same logical form as does the example problem in Table 1. "Jones fired the gun" has been inserted in place of the variable *p* in the example problem. (We might have used "Eisen-

---

[8] The term "variable" is used in quite different senses by the logician and the psychologist. If the context in which the term occurs is taken into consideration, no confusion need arise. It would be better, of course, if the same term were not used in a variety of senses.

TABLE 2
NON-FORMAL PROBLEM*

| SECTOR I | SECTOR II |
|---|---|
| 1. If Jones fired the gun or Carr was present when Smith was killed, then either Brown perjured himself or the prosecuting attorney is incompetent.<br>2. If it was the case either that Carr was present when Smith was killed or that Brown perjured himself, then the police are corrupt.<br>3. The police are not corrupt.<br>4. Jones fired the gun. | The prosecuting attorney is incompetent. |
| SECTOR III | SECTOR IV |
| 5.<br>6.<br>7.<br>etc. | |

*This particular problem can be found in Ambrose and Lazerowitz (1, p. 3). Their discussion may be helpful to the reader.

hower signed the document" or any other proposition). In a similar way other propositions have been inserted for other variables.[9] It is clear that the experimenter may vary the "content" of a problem while retaining its basic structural properties. The fact that calculus tasks are amenable to such content changes makes them particularly adaptable for attitude studies.[10]

### 3. *Complexity*

In order to solve the example problem in Table 1, the subject is required to take a series of mediating steps prior to the final step which constitutes task solution. This particular problem is a relatively complex one, since it involves a number of premises (four), the application of several rules (six), and the repeated application of the same rule (Modus Ponens, three times). Problems may be constructed with one-step, two-step, . . . , $n$-step solutions. Thus, the experimenter can generate problems of varying levels of complexity.

---

[9] When abstract Boolean algebra is interpreted as a calculus of propositions and the symbols are read in English, then "$\supset$" is read "if . . . then—," "v" is read "or" (inclusive), "$\sim$" is read "not," and "$\cdot$" is read "and."

[10] "Mode of symbolization" and "content" could be treated cogently under one heading. In a certain sense, for example, using a Chinese character or some other mark in place of a letter may be thought of as changing the "content" of the problem, and using propositions stated in English may be thought of as changing the "mode of symbolization." A distinction has been made here in order to indicate different approaches that might be taken in experimentation.

He may take "number of steps required for solution" as a simple measure of problem complexity, or utilize more subtle measures which take into consideration such factors as number of premises and number of different rules required. It is beyond the scope of this paper to treat the concept of "complexity." Nevertheless, on a preanalytic basis it seems that there are adequate means at hand for differentiating calculus problems as to level of complexity and for determining which measures of complexity are most appropriate for particular purposes.

### 4. *Progress Toward Solution*

Many of the puzzles, conundrums, and other problems used in studies of higher order problem solving make it difficult for the experimenter to record the progress of the subject toward task solution. The subject is presented with a problem, and solution is an all-or-none process from the observer's standpoint. The subject may be asked to give an introspective account of the possibilities he has considered. However, the usefulness of a number of easily differentiated check points which must be traversed if the problem is to be solved is apparent. The step-by-step character of deductive proof furnishes the experimenter with such check points. This is not to suggest that the psychological process of problem solving has any one-to-one correspondence with the logical structure of proof. To identify "laws of logic" with "laws of thought" or, in general, to confuse the logical aspects of problems and the psychological aspects of problem solving is a mistake not infrequently made. Leeper (11) and Vinacke (21) warn against such an error, and most mathematical logicians are careful to avoid this confusion.

The syllogism is one of the devices which psychologists have drawn from formal logic to use in studies of problem solving (17, 19, 24, 26). It can be used effectively and has some of the desirable characteristics of propositional calculus tasks. However, all syllogisms are of approximately the same level of complexity, and deduction is a one-step procedure. If an investigator wishes to use a logical device which, like the syllogism, involves the internal structure of propositions, then the functional calculi are excellent for this purpose. Furthermore, the employment of these calculi enables the experimenter to avoid certain difficulties found in traditional treatments of the syllogism.[11]

### 5. *Pseudo Problems*

In studies of frustration, perseveration, aspiration level, etc., the experimenter may wish to make use of some insoluble or pseudo problems; i.e., problems for which no set of steps is either necessary or sufficient for task completion or for making progress toward task completion (16). It is a simple matter to generate such pseudo tasks within the calculus. For instance, the example problem in Table 1 becomes insoluble if "∼" is placed before the *"s"* in Sector II. There is a distinct advantage in being in a position to devise pseudo problems which subjects will take seriously.

---

[11] See Ambrose and Lazerowitz (1) and Quine (18).

In making up problems, the experimenter will undoubtedly consider many more problems than he will use. In the case of extremely complex problems, it can be time-consuming to attempt to deduce the conclusion of every "candidate" problem in order to determine whether it can be solved. Fortunately, there are simple decision procedures (e.g., Quine's, 18, "full sweep") which enable the experimenter to decide whether a given conclusion can be deduced without actually doing the deduction. These procedures give a definite *yes* or *no* answer to the question "Can the problem be solved?", whereas failure to deduce a conclusion is not conclusive evidence that a problem is insoluble.

### 6. *Transfer*

One advantage of calculus tasks in their abstract form is that few subjects will have had previous specific experience with them. In addition, the working of these problems does not presuppose a knowledge of mathematics, including arithmetic and ordinary algebra. On the contrary, the calculus of propositions has been used by some logicians and mathematicians as the foundation upon which other systems have been deductively erected. It is probably true that subjects who have excelled in manipulating symbols will find calculus problems especially congenial, but no matter what kinds of experimental tasks are chosen there will be some subjects whose previous experience will aid them in approaching and working on the problems.

The material presented thus far has been placed within the context of experiments on individual problem solving. Calculus tasks are equally adaptable for use in small group problem solving studies. The latter application has been treated in some detail by Moore, Rudner, and Anderson (16).

## D. Development of Instructions[12]

Although the calculus of propositions appeared to present a number of advantages as a source of experimental tasks, there was little encouraging evidence to indicate that subjects could be easily trained and motivated to work on the tasks. The methods of teaching symbolic logic employed in most universities seemed much too involved and time-consuming for the training of experimental subjects.

---

[12] R. S. Rudner, Tufts College, worked with the authors in the development of the tasks and instructions.

It was decided that an adequate set of instructions for experimental use should meet the following criteria:

1. The instructional period should be brief, no more than one to two hours in length.
2. The instructions should be standardized, such that instructor or teaching method would not represent a major variable in determining performance level of subjects.
3. The instructions should be easily administrable.
4. There should be enough rules that in any particular experiment *no more* than the pilot set would be required.
5. The task should be interesting to subjects.

Rather than develop a complete set of instructions by the "arm-chair" method and later test them on subjects, it was decided to include subjects from the very beginning of the development phase. It was felt that their questions and difficulties would aid in the choice of instructional techniques.

A total of 50 subjects, divided into 10 groups, were brought into the laboratory. They were Naval enlisted men, with a mean *GCT* score of 2.4 (range 1-4) and a mean educational level of 11.5 years (range 8-16 years).

Teaching methods evolved gradually from informal lecture with blackboard demonstration and student notes to an apparently effective set of techniques. The following items were developed:

1. A recorded lecture, allowing for subject participation but not requiring additional comment by the experimenter.
2. Individual magnetic boards with sets of standard symbols for subjects to manipulate in accordance with instructions from the recording.
3. Individual sets of 12 plastic cards, showing examples of applications of rules.
4. Printed individual rule sheets.
5. A wall chart containing an example problem explained in the recording.

The entire procedure was represented as a problem in handling coded messages received in an intelligence center. Thus the task consisted of finding a "hidden" or "concealed" specific piece of information within a total coded message. There was nothing in the instructions to suggest that problems were drawn from mathematical logic or that subjects were involved in making deductions. The "coding" device seemed to have considerable interest value for subjects.

Twelve rules were retained as follows: Commutativity, Contraposition, Iteration, Associativity, DeMorgan Laws, Definition of Conditional, Distribution, Detachment, Addition, Adjunction, Modus Ponens, and Transitivity of Conditional. The rules were presented as rules of the code and were identified to the subjects only by number. Thus as in Sector IV, Table 1, subjects recorded only numbers of rules and steps to which those rules were applied; e.g., from Step 4 by Rule 9.

Following the instruction period, all subjects were given test sets of problems. Forty of the 50 subjects were able to work these problems with sufficient proficiency that they would have been ready to participate in experiments using calculus tasks. The total time spent in instructions, of course, varied for the different groups. In the last four groups this time had been reduced to one hour and 20 minutes and 18 out of 20 of these subjects performed satisfactorily on a set of test problems.[13]

The results of the instruction development period clearly indicate that subjects can be easily and rather quickly taught to work on the tasks. As has been suggested earlier, other experimenters might prefer different instructional methods or frames of reference, but it is believed that the content and teaching methods which proved efficient in the present situation can be easily adapted for other uses.

## E. SUMMARY

The calculus of propositions has been adapted for use as a source of tasks for human problem solving experiments. This contribution of modern logic is differentiated from other applications of logic to the empirical sciences; i.e., axiomatizations, explications, and tests of the soundness of deductions.

The calculus tasks possess the following characteristics which make them efficacious for experiments in the area of higher order problem solving: (*a*) mode of symbolization and content may be varied, (*b*) problems of varying degrees of complexity may be generated, (*c*) subjects' progress toward solution may, in part, be followed by the experimenter, (*d*) insoluble problems (pseudo tasks) may be devised within the same symbolic context as soluble problems, (*e*) simple decision procedures are available to the experimenter, (*f*) few subjects will have had any previous experience with tasks of this type, (*g*) tasks are adaptable for group as well as individual experiments, and (*h*) subjects may be easily trained and motivated to work on the tasks.

---

[13] By "satisfactory performance" we mean that a subject can correctly apply rules, not necessarily solve highly complex problems.

## REFERENCES

1. AMBROSE, A., & LAZEROWITZ, M. Fundamentals of Symbolic Logic. New York: Rinehart, 1948.
2. BLOCH, B. A set of postulates for phonemic analysis. *Language*, 1948, **24**, 3-36.
3. BLOOMFIELD, L. A set of postulates for the science of language. *Language*, 1926, **2**, 153-164.
4. COOLEY, J. C. A Primer of Formal Logic. New York: Macmillan, 1947.
5. FITCH, F. B. Symbolic Logic. New York: Ronald, 1952.
6. GREENBERG, J. E. The logical analysis of kinship. *Philos. Sci.*, 1949, **16**, 58-64.
7. HEMPEL, C. G. Fundamentals of concept formation in empirical science. *Internat. Ency. Unified Sci.*, Vol. II, No. 7. Chicago: Univ. Chicago Press, 1952.
8. HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. B. Mathematico-deductive Theory of Rote Learning. New Haven: Yale Univ. Press, 1940.
9. HUMPHREY, G. Thinking. New York: Wiley, 1951.
10. JOHNSON, D. M. A modern account of problem solving. *Psychol. Bull.*, 1944, **41**, 201-229.
11. LEEPER, E. Cognitive processes. In Stevens, S. S. (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951. (Pp. 730-757.)
12. MCCULLOCH, W. S., & PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, 1943, **5**, 115-133.
13. MCKINSEY, J. C. C., SUGAR, A. C., & SUPPES, P. Axiomatic foundations of classical particle mechanics. Report on work under Contract N-onr-225 (13), 1952.
14. MOORE, O. K. Nominal definitions of "culture." *Philos. Sci.*, 1952, **19**, 245-256.
15. MOORE, O. K., & LEWIS, D. J. Learning theory and culture. *Psychol. Rev.*, 1952, **59**, 380-388.
16. MOORE, O. K., RUDNER, R. S., & ANDERSON, S. Logic, computers, and small group research. (Unpublished manuscript.)
17. MORGAN, J. J. B., & MORTON, J. T. The distortion of syllogistic reasoning produced by personal conviction. *J. Soc. Psychol.*, 1944, **20**, 39-59.
18. QUINE, W. V. O. Methods of Logic. New York: Holt, 1950.
19. SELLS, S. B. The atmosphere effect: an experimental study of reasoning. *Arch. Psychol.*, 1936, No. 200.
20. UNDERWOOD, B. J. Experimental Psychology. New York: Appleton-Century-Crofts, 1949.
21. VINACKE, W. E. The Psychology of Thinking. New York: McGraw-Hill, 1952.
22. VON NEUMANN, J., & MORGENSTERN, O. Theory of Games and Economic Behavior. Princeton: Princeton Univ. Press, 1947.
23. WERTHEIMER, M. Productive Thinking. New York: Harper, 1945.
24. WILKINS, N. C. The effect of changed material on ability to do formal syllogistic reasoning. *Arch. of Psychol.*, 1928, No. 102.
25. WOODGER, J. H. The Axiomatic Method in Biology. Cambridge: Cambridge Univ. Press, 1937.
26. WOODWORTH, R. S., & SELLS, S. B. An atmosphere effect in formal syllogistic reasoning. *J. Exp. Psychol.*, 1935, **18**, 451-460.

*Naval Research Laboratory*
*Systems Division*
*Washington 25, D. C.*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 161-175.

# JUDGMENT OF VISUAL AND POSTURAL VERTICAL AFTER BRAIN INJURY*†

*Psychophysiological Laboratory, Department of Psychiatry and Neurology, New York University—Bellevue Medical Center*

HANS-LUKAS TEUBER AND MORTIMER MISHKIN

## A. THE PROBLEM

Lesions in the frontal lobes anterior to the motor cortex have been claimed to produce a great variety of deficits in man and lower primates. The disturbances reported range from inadequacy on complex problem solving tasks to impairment of simple sensorimotor skills. Few of these claims have remained without contradiction. The different forms of breakdown assumed to occur in "higher" functions (8, 16, 18) are hard to demonstrate after restricted damage (4, 20, 26), and are difficult to distinguish from some non-specific deterioration of behavior in those cases where breakdown does occur (10). Moreover, it does not seem established either for monkey (6, 19) or man (35, 37) that frontal lesions produce more severe deficits of higher functions than do post-central lesions.

Claims of disturbance in simple sensorimotor functions are somewhat more specific but no less controversial. Two types of deficit have been proposed: a postural and a visual one. The idea of a change in postural control (9) dates back to Munk's observation of "trunk ataxia" as a temporary sequel of bifrontal ablations in macaques (30, 31), and to subsequent clinical reports on tonus-disturbances in patients with frontal tumor (5), vascular accident (15, 17), or gunshot wound (13). Critical surveys, such as Feuchtwanger's (11), have shown that these symptoms are infrequent and transient.

Visual disturbances after frontal lesions are suggested by the peculiar "pseudo-hemianopsia" which Loeb first noted in dogs (23), and which Bianchi (3) and others (7, 21, 22) observed in monkeys, with unilateral frontal lesions. Yet most forms of visual discrimination are relatively intact after frontal ablations in the monkey (6, 19). Recent work with a variety of visual tests, in man, has likewise failed to yield any specifically frontal deficits (35, 37).

---

*Received in the Editorial Office on April 1, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

†The program of research of which this study is a part is supported by the Commonwealth Fund of New York, and by the Office of the Surgeon General (Department of the Army), Contract DA-49-007-MD-312.

Thus, neither purely postural nor purely visual disturbances seem likely hypotheses for a further exploration of effects of frontal lobe injury. But the persistent reports of at least transient difficulties in vision or posture suggest that special conditions of testing might reveal more general or more permanent deficits in these areas. One possibility is that frontal lesions might produce abnormal interaction between vision and posture.

A classical situation involving these two modalities is provided by Aubert's experiment in which the subject, in a dark room, adjusts a luminous line to the vertical (1). As long as the subject is upright, his settings are quite accurate (12, 38). When the subject's body is tilted, however, his settings deviate from the true vertical, the extent and direction of the error depending in part on the angle and direction of the body tilt (24, 29). There are large individual differences in this phenomenon among normal subjects (14, 42). Nevertheless, men with cerebral lesions have been reported to make significantly greater errors than normal controls (2).

The present study was undertaken to see whether men with frontal involvement show larger deviations in the Aubert situation than men with lesions elsewhere in the cerebrum. Moreover, we hoped to determine, if such a selective relationship were established, whether the impairment was related to the particular intermodal aspects of the task (postural tilt vs. visual orientation), or whether alternative interpretations could be entertained. Some of the obvious alternatives are: (*a*) that a deficit on the Aubert task is due, after all, to purely postural, or purely visual difficulties, or, (*b*) that it is due to some abnormal interference between stimulus events, but that different modalities need not be involved; in that case, interfering sense data within a single modality might yield equally aberrant responses in the frontal group.

Four experiments were set up accordingly: one to answer the main question (whether frontal cases act abnormally in the Aubert situation, Expt. 2 below), and three supplementary experiments designed to deal with alternatives to the interpretation that deficits in the Aubert situation result from abnormal intermodal conflict. Schematically, the supplementary experiments consisted of a "purely visual" task, a "purely postural" task, and a "visual-visual interference" situation.

Concretely, the four tasks were (in the order given):

Expt. 1. Setting a luminous line to vertical, in a dark room, with body upright (visual task).

Expt. 2. Setting a luminous line to vertical, in a dark room, with body tilted (visual-postural task).

Expt. 3. Setting a line to vertical, with body upright, the line appearing against an obliquely striped visual background (visual-visual task).

Expt. 4. Setting the body vertical, with vision excluded, after various periods of body tilt (postural task).

## B. Subjects

The 60 subjects (40 patients and 20 controls) for these experiments were drawn from the population accessible to this laboratory (35): viz., 185 male veterans with proven loss of brain substance due to penetrating gunshot wounds (experimentals), and 101 with gunshot wounds of peripheral nerves, but without CNS involvement (controls). The basis upon which subjects were selected for the present experiments was (*a*) presence of a lesion in the anterior third, but not in the posterior third of the brain, in 20 cases (anterior group, *A*); (*b*) presence of penetration in the posterior third of the brain, but not in the anterior, in 20 cases (posterior group, *P*); (*c*) presence of a penetrating injury in one extremity, but not the central nervous system, in 20 cases (control group, *C*). In addition, some care was exercised in testing approximately equal numbers of cases with injury involving the right and left sides. The resulting distribution of unilateral and bilateral cases is shown in Table 1, together with the median age and range of ages for each of the three groups.

It should be noted that both anterior and posterior groups in this study

TABLE 1

Distribution of Subjects by Site of Injury and by Age

| | Right | Left | Bilateral | Total |
|---|---|---|---|---|
| *Brain Injury* | | | | |
| Anterior Third | 3 | 6 | 5 | 14 |
| Anterior and Intermediate Thirds | 3 | 0 | 3 | 6 |
| Total "Anterior" Group | 6 | 6 | 8 | 20 |
| Median Age: 31. Range: 26 to 44. | | | | |
| Posterior Third | 5 | 6 | 2 | 13 |
| Posterior and Intermediate Thirds | 2 | 2 | 3 | 7 |
| Total "Posterior" Group | 7 | 8 | 5 | 20 |
| Median Age: 31. Range: 28 to 43. | | | | |
| Total Brain-Injured | 13 | 14 | 13 | 40 |
| *Nerve Injury* (Controls) | | | | |
| Arm | 6 | 5 | — | 11 |
| Leg | 5 | 4 | — | 9 |
| Total Controls | 11 | 9 | — | 20 |
| Median Age: 31. Range: 27 to 42. | | | | |

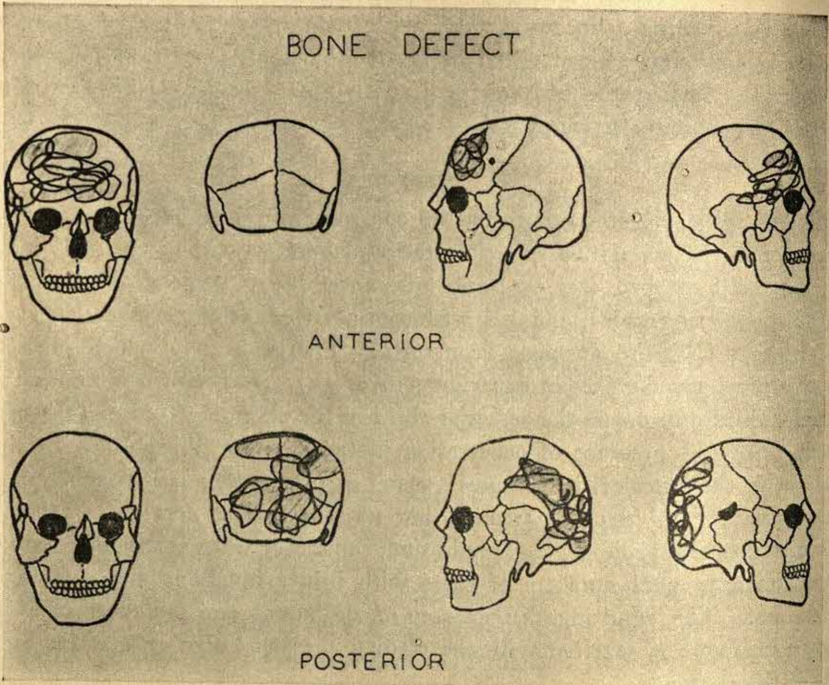


FIGURE 1
COMPOSITE DIAGRAMS SHOWING LOCATIONS OF BONE DEFECT IN 20 CASES OF PENETRATION OF THE ANTERIOR BRAIN SUBSTANCE (UPPER ROW) AND IN 20 CASES OF PENETRATION OF THE POSTERIOR BRAIN SUBSTANCE (LOWER ROW): RETAINED FOREIGN BODIES WITHIN THE CALVARIUM ARE SHOWN IN SOLID BLACK

included some cases of lesions overlapping into the intermediate third of the brain (see Table 1). The location of bone defects in the 40 cases of brain injury is shown in Figure 1. All men (patients and controls) were veterans of the second World War who had sustained their injuries 7-10 years prior to testing. Post-traumatic epilepsy of one form or another existed in six men in the *A* group, and in seven men in the *P* group. Of these 13 patients, five in each group (*A* and *P*) received combined barbiturate and dilantin medication which had succeeded in suppressing seizures and seizure-equivalents in three men in each group.

## C. APPARATUS

A tilting chair[1] was used throughout these experiments (Figure 2a). It provided foot, arm, and head rests and could be set vertically, or inclined, in *S*'s frontal plane, 28° to the left or right. In addition to the tilting chair, two visual displays were employed, one in Experiments 1 and 2, the other in Experiment 3.

---

[1] We are indebted to Dr. H. A. Witkin for providing us with a duplicate of the tilting chair employed in his studies of the Aubert phenomenon and related effects in normals (41).

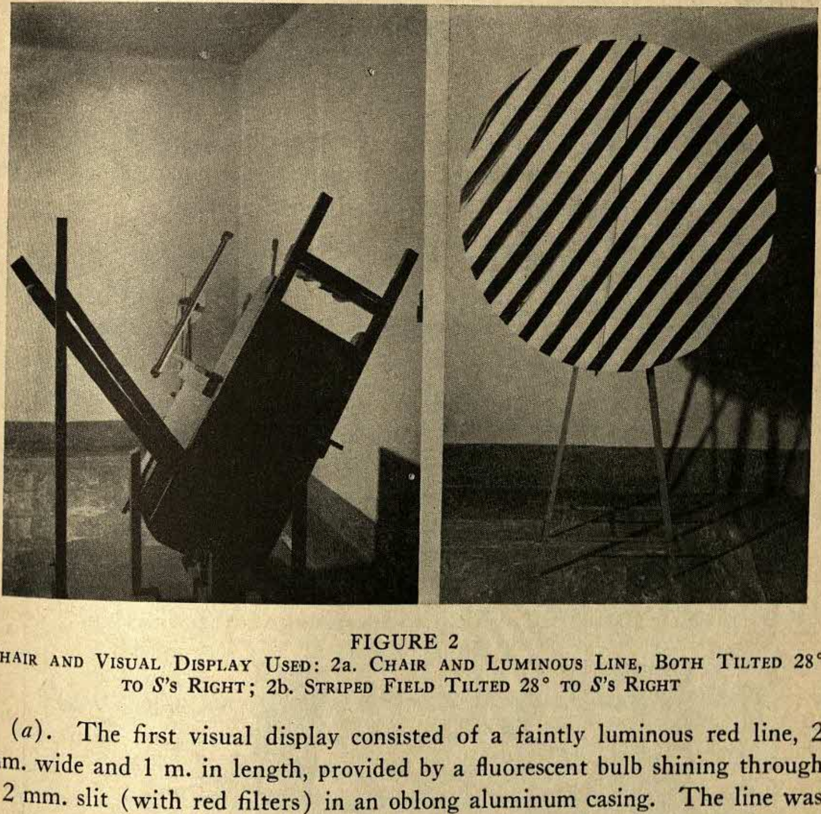


FIGURE 2
CHAIR AND VISUAL DISPLAY USED: 2a. CHAIR AND LUMINOUS LINE, BOTH TILTED 28° TO *S*'s RIGHT; 2b. STRIPED FIELD TILTED 28° TO *S*'s RIGHT

(*a*). The first visual display consisted of a faintly luminous red line, 2 mm. wide and 1 m. in length, provided by a fluorescent bulb shining through a 2 mm. slit (with red filters) in an oblong aluminum casing. The line was mounted at its midpoint and could be set at any angle in the *S*'s fronto-parallel plane (see Figure 2a).

(*b*). The second visual display consisted of a metal disc, 1 m. in diameter, which was covered with alternating black and white stripes, 5 cm. in width. The disc was mounted at its midpoint in the *S*'s fronto-parallel plane. One cm. in front of the disc and bisecting it, a black thread was mounted so that it could be set at any angle independently of the angle of the stripes. A cardboard tunnel restricted *S*'s vision to the thread-and-disc-display (Figure 2b).

Graduated circular scales on each apparatus permitted *E* to read off the positions of the subjective vertical (luminous line, or black thread) as set by *S*. Zero degrees designated true verticality. Deviations from the vertical were expressed as being "to the left," when the top of the line (or thread) was inclined to *S*'s left, and "to the right," when the top was inclined to the *S*'s right.

### D. Procedures

Subjects were tested either in a single one-hour session, or in two half-hour sessions held on separate days. The sequence of tests was the same throughout. All 60 *S*'s participated in Experiments 1 and 2. Fifteen in each of the three groups continued to complete Experiments 3 and 4. The four experiments are illustrated schematically in Figure 3.

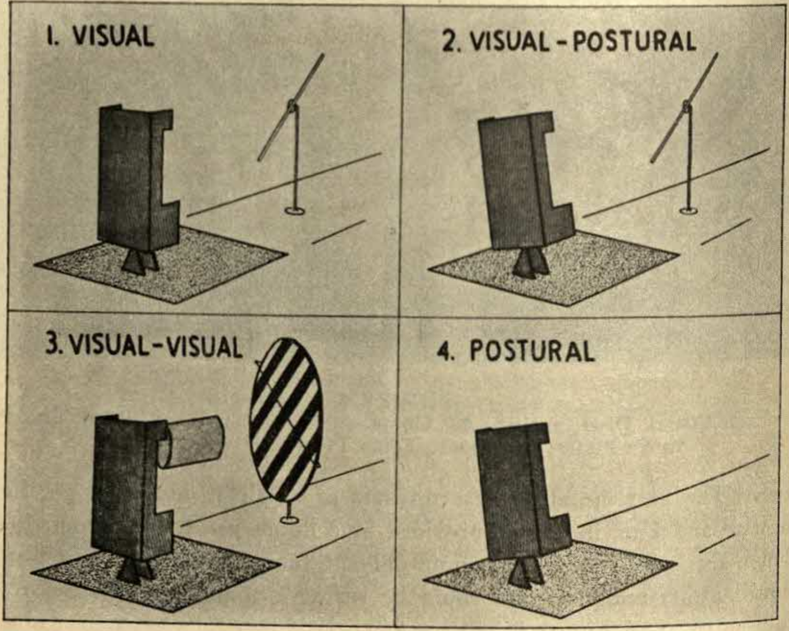


FIGURE 3
Schematic Diagrams Showing the Four Experiments in the Order in Which They Were Performed

#### 1. *Experiment 1: Judging Visual Vertical With Body Upright*

After entering the experimental room, *S* was seated with the tilting chair in the upright position, and the head support adjusted. The luminous-line apparatus was placed 2 m. from the *S*, with the midpoint of the line at *S*'s eye level. The *S* was blindfolded with opaque goggles which he was instructed to wear over his eyes between trials, and to shift to his forehead during trials. The room was darkened, and the line (the only visible object in the room) was presented to *S*'s view, with the top of the line 28° to the *S*'s left. *E* moved the line slowly towards the vertical (approximately 3°/sec), and *S* told him when he thought that the line was truly vertical. Reversals in direction of movement were permitted if *S* reported that the vertical had been passed. The position of the line was recorded while the *S*'s eyes were blindfolded. A total of six trials, in which the initial position of the line was alternately 28° to the *S*'s left and right, was run in this manner, all with the subject upright in the chair.

### 2. *Experiment 2: Judging Visual Vertical With Body Tilted*

The chair and subject were tilted 28° to the *S*'s left, and the luminous-line apparatus was shifted so that the center of the line was again directly in front of *S*'s eyes. The six trials were then given as before, with the initial position of the line alternately 28° left and right. A final series of six such trials was then given with chair and subject tilted 28° to the *S*'s right.

The blindfolds were removed, the room lit, and the *S* given a rest period during which he walked about the room. In those cases where the testing was divided into two sessions, the break occurred at this point. Testing was then resumed with Experiment 3.

### 3. *Experiment 3: Judging Visual Vertical Against Tilted Visual Background*

The thread-and-striped disc display was placed 2 m. in front of *S*, who was again seated upright in the tilting chair. The cardboard tube was placed around *S*'s head and adjusted so that it excluded all other objects from view, but permitted maximal view of the display. The *S* was instructed to keep his eyes closed between trials, and to open them only during trials. *E* then set the striped disc and the thread so that the tops of both stripes and thread were 28° to the *S*'s left. *S* opened his eyes and *E* rotated the thread slowly (approximately 3°/sec) until S indicated that it appeared vertical. The position of the thread was recorded. Six trials were given in this manner, with the stripes 28° to the left, and with the initial position of the thread alternately to the *S*'s left and right. The striped field was then set with the top of the stripes 28° to the *S*'s right, and another series of six trials given, with the initial position of the thread alternating as before.

### 4. *Experiment 4: Judging Postural Vertical With Vision Excluded*

In the final procedure, *S* was blindfolded and seated upright in the tilting chair. The chair was inclined 28° to the *S*'s left. At the end of two minutes, *E* brought the chair slowly toward the vertical (approximately 1°/sec) until *S* indicated that he felt upright. Immediately following the judgment the chair was brought, without stopping at the objective vertical, back to its 28° tilt. Four more judgments were obtained in this manner, after successive one-minute tilts, making a total of five judgments. Following a five-minute rest period during which *S* walked about the room, the entire procedure (five judgments) was repeated with the chair tilted 28° to the *S*'s right.

## E. Results

### 1. *Experiments 1 and 2*

The data from Experiments 1 and 2 are presented together. Scores were expressed as errors (deviations of the top of the line in degrees of arc) to the left or to the right (from *S*) of the objective vertical. A score for each individual *S* on each condition (three trials per condition) was computed by summing, separately, those errors made to one side and those made to the other side of the objective vertical, and then subtracting the smaller from the larger value. The average of this value, termed the average constant error, expresses a tendency for the *S*'s judgments to deviate from the true vertical in a consistent direction. The means and standard deviations of the constant error for each injury group, in each of the six combinations of body position and line position, are given in Table 2.

TABLE 2
Means and Standard Deviations (in Parenthesis) for the Constant Error in Degrees, Experiments 1 and 2 ($N = 20$)

| Body position | Vertical | | Left | | Right | |
|---|---|---|---|---|---|---|
| Line position<br>Injury groups | Left | Right | Left | Right | Left | Right |
| Anterior | 0.2L | 0.2R | 2.6R | 7.4R | 9.1L | 5.2L |
| | (1.2) | (1.2) | (4.7) | (5.1) | (4.9) | (4.2) |
| Posterior | 0.6L | 0.2R | 0.5R | 5.0R | 5.8L | 1.8L |
| | (1.8) | (1.7) | (6.8) | (7.2) | (7.3) | (6.9) |
| Control | 0.5L | 0.1L | 0.4R | 3.2R | 3.0L | 0.5L |
| | (1.1) | (0.3) | (3.6) | (4.5) | (4.3) | (3.0) |

With body upright (Experiment 1) all groups made only small errors in setting the luminous line; none of the group differences, as tested by the method of ranking unpaired replicates (40) were significant (see Table 3).

TABLE 3
DIFFERENCES BETWEEN MEANS IN DEGREES AND *P* VALUES OF RANK DIFFERENCES FOR CONSTANT ERRORS OF INJURY GROUPS, EXPERIMENTS 1 AND 2

| Body position | Vertical | | Left | | Right | |
|---|---|---|---|---|---|---|
| Line position<br>Comparison of injury group | Left | Right | Left | Right | Left | Right |
| A-C | —.3 | .3 | 2.2* | 4.2** | 6.1** | 4.7** |
| A-P | —.4 | .0 | 2.1** | 2.4* | 3.3* | 3.4** |
| P-C | .1 | .3 | .1 | 1.8 | 2.8 | 1.3 |

A = Anterior; C = control: P = Posterior.
*Rank differences sig. at .05 level of confidence.
**Rank differences sig. at .01 level of confidence.

With body tilted (Experiment 2) the means of the average constant errors are larger for all groups. Inspection of Table 2 shows that the errors vary systematically with the direction in which *S*'s body was tilted: with the body tilted right the line was placed by all groups predominantly to the left of true vertical, and conversely for body tilted left (*E*-phenomenon, 29, 42). In addition, adjustment of the line from an inclination opposite to *S*'s body tilt led to greater constant errors than adjustment of the line from inclination to the same side as the body tilt (39). However, the injury groups differed in the extent to which they were affected by these conditions. Under all conditions of body tilt the average constant errors are greatest in the *A* group, less in the *P* group, and least in the controls. The differences between the anteriors and posteriors and between the anteriors and controls are all significant at accepted levels of confidence (see Table 3). The differences between the posteriors and controls, though all in a direction indicating greater errors for the posteriors, do not attain significance.

It appears, then, that abnormally large constant errors in the visual-postural task are associated with anterior brain injury. The question arises whether these constant errors might not reflect the influence of some other factor more directly related to location of the lesion. The laboratory records of each of the 20 men in the *A* and *P* groups were searched for possible symptom associations, but this search was negative. Extent of error in the visual-postural task was unrelated to presence of motor deficits (including signs of pyramidal tract involvement), sensory defects, traumatic epilepsy, aphasia, or visual field defects. In fact, 11 men in the anterior group were entirely free from such neurological symptoms at the time of testing.

A further question is whether the laterality of brain injury (right, left, or bilateral) influenced the finding of large errors for the anterior cases in Experiment 2. Reference to Table 1 indicates that approximately equal numbers of right, left, and bilateral cases were included in each of our brain-injured groups. Thus, any consistent effects of laterality would presumably be equally distributed among the groups.

To determine the effects of laterality as an independent factor, the total group of 40 brain-injured subjects was divided into three equal subgroups—left, right, and bilateral—with 13 in each (one left posterior case—see Table 1—was excluded). The scores of the right, left, and bilateral groups and of a group of 13 randomly selected controls were then compared by Wilcoxon's non-parametric method (40). This test revealed only two differences which attained significance (both at the .01 level): those between right-hemisphere cases and controls for the two conditions of body tilted right. (The means for the right-hemisphere group were 8.8°L for the line tilted left condition and 5.6°L for the line tilted right condition; the comparable means of the 13 controls were 2.2°L and 0.1°L, respectively.) Since no significant differences could be demonstrated between the right-hemisphere and bilateral groups, or between the right-hemisphere and left-hemisphere groups, the importance to be attached to the positive findings (comparing the right-hemisphere group and controls) cannot be easily evaluated.

### 2. *Experiment 3*

The means and standard deviations of the constant errors for Experiment 3 are given in Table 4. The scores show a consistent tendency for the *S*'s in all groups to err in the direction of the stripes. The data indicate, further, that the posterior group made larger errors than either the controls or the anterior group in each of the four conditions. In direction, at least, the result of this comparison between anteriors and posteriors is the reverse of that found in Experiment 2. The differences between the constant errors are small, however, and none of them attain significance by the method of ranking unpaired replicates. An analysis in terms of laterality likewise yielded no reliable differences.

TABLE 4
MEANS AND STANDARD DEVIATIONS (IN PARENTHESIS) FOR THE CONSTANT ERROR IN DEGREES, EXPERIMENT 3 ($N = 15$)

| Stripe position | Left | | Right | |
|---|---|---|---|---|
| Thread position | Left | Right | Left | Right |
| Injury groups | | | | |
| Anterior | 0.8L<br>(2.0) | 0.5L<br>(2.3) | 1.5R<br>(2.0) | 2.1R<br>(2.4) |
| Posterior | 1.5L<br>(2.8) | 1.9L<br>(4.0) | 1.6R<br>(4.6) | 2.1R<br>(4.7) |
| Controls | 1.4L<br>(2.3) | 0.9L<br>(2.0) | 0.3R<br>(1.0) | 0.8R<br>(1.2) |

### 3. *Experiment 4*

In Experiment 4, where subjects adjusted their body to the subjective vertical after various periods of body tilt, all *S*'s erred, with rare exceptions, by falling short of the true vertical (see Table 5). Thus, after body tilts to the left they stayed too far to the left, and after body tilts to the right they stayed too far to the right. Analysis of the constant errors revealed no significant trends with respect to locus of injury. The "purely postural" task (Experiment 4) thus did not differentiate our experimental groups.

TABLE 5
MEANS AND STANDARD DEVIATIONS (IN PARENTHESIS) FOR THE CONSTANT ERROR IN DEGREES, EXPERIMENT 4 ($N = 15$)

| Body position<br>Injury group | Left | Right |
|---|---|---|
| Anterior | 6.8L (4.0) | 5.8R (3.2) |
| Posterior | 7.0L (3.5) | 6.6R (4.3) |
| Control | 8.0L (3.0) | 7.1R (3.1) |

## F. SUMMARY AND DISCUSSION

The outcome of this series of experiments may be summarized briefly as follows: The anterior group made greater errors than did either the normals or the posterior group in setting the visual vertical under conditions of abnormal posture (Experiment 2). The posterior group, on the other hand, did not differ significantly from the controls.

In the situation where the visual vertical had to be set against a conflicting visual background (Experiment 3) the difference between the brain-injured groups was reversed, the posterior group performing more poorly than either the anterior group or the controls. Although these latter differences were not reliable, the reversal in direction makes it unlikely that the poor performance of the anterior cases in the Aubert situation is related to perceptual "conflict," per se. Finally, neither the purely visual task (Experiment 1) nor the purely postural task (Experiment 4), produced differential results. It appears at present, therefore, that the inferior performance of the anterior group in the Aubert situation is related to the particular interaction between visual orientation and abnormal posture.

Although the deficit is clearly related to anterior locus of injury, it is not apparent which part of the anterior substance is involved. The deficit might be produced by damage to frontal granular, or frontal agranular cortex, or by damage to anterior basal ganglia (25, 33). The lack of neurological symptoms in many of the cases showing deficit on the Aubert task makes it appear that extensive involvement of either agranular cortex or basal ganglia is not a necessary condition. These considerations make it likely that prefrontal (granular) cortical damage may be sufficient to produce the symptom. Questions of detailed localization aside, the salient finding is the appearance of deficit on the Aubert task, in cases which are free from any demonstrable clinical symptoms of cerebral damage.

Performance on the Aubert task by brain-injured cases was made the subject of an earlier investigation by Manfred Bender and Richard Jung (2). These authors report that men with gunshot wounds of "frontal, rolandic and parietal regions" made substantially larger errors than did men with temporal or occipital injury. In contrast to the present experiment the earlier study established its group differences in the settings of the visual vertical, not only with body tilted, but with body upright. (In the latter situation, the errors of Bender and Jungs' patients consisted of a tendency to displace the luminous line in a direction opposite to the side of the lesion.)

A possible source of the discrepancy between the results of the two studies might lie in the differences between the populations studied. Most of the data for the earlier study were gathered on German casualties during the war years 1940-42; our cases were studied 7-10 years after wounding. It is conceivable that the shorter interval between dates of injury and of testing in the earlier study was responsible for the finding of abnormalities in the body-upright condition. Aberrant performance under conditions of body tilt might then represent a stable residual of an initially more generalized disturbance. Similarly, the recency of the injuries in the earlier study may have contributed to the finding of impairment in a group with a somewhat more widespread cerebral involvement. For the patients with older injuries reported here, results were negative for the posterior (including presumably parietal) cases.

If these considerations are correct, the present study, together with previous reports from this laboratory, provide evidence suggesting a double dissociation of symptoms for men with penetrating gunshot wounds of the skull. On certain complex visual tasks men with posterior injury have been found to perform more poorly than men with anterior injury (35, 37). On the Aubert task the converse appears to be true.

Such double dissociation of symptoms has already been well established for animals. Macaques with posterior (particularly temporal) lesions show considerably greater difficulties on certain visual discriminations than do animals with anterior frontal lesions (6, 19, 27). Conversely, macaques with frontal ablations exhibit nearly complete loss on delayed response tests, whereas animals with posterior lesions are essentially unimpaired (19).

At present, it is of course difficult to see even a superficial similarity between what might constitute the delayed-response problem for monkeys, and the peculiar perceptual interaction involved for man in the Aubert task. It is conceivable, however, that the common feature of these disorders is a derangement in the integration of postural with visual (as well as with other exteroceptive) functions (28). From this point of view other symptoms which have been reported following frontal damage, such as forced circling, hypermotility, or "pseudo-hemianopsia" in monkey (21, 22), and impaired visual searching performance in man (33, 34, 36) may also be related to a disorganization of postural and exteroceptive mechanisms. In any event it would seem that an elucidation of frontal lobe functions, for animal and man, is more likely to be obtained by exploring these postural-visual and related tasks, than by continuing the search for supposed disturbances in "higher intellectual" processes.

## REFERENCES

1. Aubert, H. Eine scheinbar bestehende Drehung von Objekten bei Neigung des Kopfes nach rechts und links. *Arch. Pathol. Anat. Physiol. (Virchow's Arch.)*, 1861, **20**, 381-383.
2. Bender, M., & Jung, R. Abweichungen der subjektiven optischen Vertikalen und Horizontalen bei Gesunden und Hirnverletzten. *Arch. f. Psychiatr. u. Z. Neurol.*, 1948, **181**, 193-212.
3. Bianchi, L. The functions of the frontal lobes. *Brain*, 1895, **18**, 497-522.
4. Blum, R. A. The effects of bilateral removal of the prefrontal granular cortex on delayed response performance and emotionality in chimpanzee. *Amer. Psychol.*, 1948, **3**, 237-238.
5. Bruns, L. Über Störungen des Gleichgewichtes bei Stirnhirntumor. *Dtsch. med. Wochenschr.*, 1892, **18**, 138-140.
6. Chow, K. L. Further studies on selective ablation of associative cortex in relation to visually mediated behavior. *J. Comp. & Physiol. Psychol.*, 1952, **45**, 109-118.
7. Clark, G., & Lashley, K. S. Visual disturbances following frontal ablations in the monkey. *Anat. Rec.*, 1947, **97**, 326.
8. Crawford, M. P., Fulton, J. F., Jacobsen, C. F., & Wolfe, J. B. Frontal lobe ablation in chimpanzee: a résumé of "Becky" and "Lucy." *Res. Publ. Ass. Nerv. & Ment. Dis.*, 1948, **27**, 3-58.
9. Delmas-Marsalat, P. Lobe frontal et équilibre. *Encéphale*, 1936, **31**, 15-91.
10. Denny-Brown, D. The frontal lobes and their function. Chap. 2, pp. 13-89, in: Feiling, A., *Ed.*, *Modern Trends in Neurology*. New York: Hoeber-Harper, 1951.
11. Feuchtwanger, E. *Die Funktionen des Stirnhirns. Ihre Pathologie und Psychologie.* (Monogr. a. d. Gesgeb. d. Neurol. u. Psychiat., *Ed.* O. Foerster u. K. Wilmanns, 1923, H. 38). Berlin: Springer, 1923. Pp. IV & 194.
12. Fischer, M. H. Die Orientierung im Raume bei Wirbeltieren und beim Menschen. *Hdb. Norm. & Pathol. Physiol.*, 1931, **15** (2nd half), 909-1022.
13. Gerstmann, J. Zur Kenntnis der Störungen des Körpergleichgewichtes nach Schussverletzungen des Stirnhirns. *Mschr. Psychiatr. Neurol.*, 1916, **40**, 354-377.

14. GIBSON, J. J. The relation between visual and postural determinants of the phenomenal vertical. *Psychol. Rev.*, 1952, **59**, 370-375.

15. GOLDSTEIN, K. Über Zeigen und Greifen. *Nervenarzt,* 1931, **4**, 453-466.

16. ———. The mental changes due to frontal lobe damage. *J. of Psychol.*, 1944, **17**, 187-208.

17. HALPERN, L. The syndrome of sensorimotor induction in disturbed equilibrium. *Arch. Neurol. & Psychiat.*, 1949, **62**, 330-354.

18. HALSTEAD, W. Brain and Intelligence. A Quantitative Study of the Frontal Lobes. Chicago: Univ. Chicago Press, 1947. Pp. XIII & 206.

19. HARLOW, H. F., DAVIS, R. T., SETTLAGE, P. H., & MEYER, D. R. Analysis of frontal and posterior association syndromes in brain-damaged monkeys. *J. Comp. & Physiol. Psychol.*, 1952, **45**, 419-429.

20. HEBB, D. O. Man's frontal lobes: a critical review. *Arch. Neurol. & Psychiat.*, 1945, **54**, 10-24.

21. KENNARD, M. A. Alterations in response to visual stimuli following lesions in the frontal lobes in monkeys. *Arch. Neurol. & Psychiat.*, 1939, **41**, 1153-1165.

22. KENNARD, M. A., & ECTORS, L. Forced circling in monkeys following lesions of the frontal lobes. *J. Neurophysiol.*, 1938, **1**, 45-54.

23. LOEB, J. Die Sehstörungen nach Verletzung der Grosshirnrinde. Nach Versuchen am Hunde. *Pflüger's Arch. ges. Physiol.*, 1884, **34**, 67-172.

24. MANN, C. W., BERTHELOT-BERRY, N. H., & DAUTERIVE, H. J., JR. The perception of the vertical: I. Visual and non-labyrinthine cues. *J. Exp. Psychol.*, 1949, **39**, 538-547.

25. METTLER, F. A. Corticofugal fiber connections of the cortex of Macaca mulatta: The frontal region. *J. Comp. Neurol.*, 1935, **61**, 509-542.

26. METTLER, F. A. (*Ed.*). Columbia-Greystone Associates: *Problems of the Human Brain: I. Selective Partial Ablations of the Frontal Cortex.* New York: Hoeber, 1949.

27. MISHKIN, M. Visual discrimination performance following partial ablations of the temporal lobe: II. Ventral surface vs. Hippocampus. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 187-193.

28. MONAKOW, C. v. Der rote Kern, die Haube und die Regio hypothalamica bei einigen Säugetieren und beim Menschen. II. Pathologisch-anatomische Untersuchungen am Menschen. *Arbeiten a. d. hirnanat. Inst. Zürich,* H. IV, 1910, 103-226 (see, esp., p. 209).

29. MÜLLER, G. E. Über das Aubertsche Phänomen. *Z. f. Psychol.*, 1916, II, **49**, 109-244.

30. MUNK, H. Über die Functionen der Grosshirnrinde. Vierte Mittheilung: Fühlsphären von Hund und Affe. *Verhdl. Physiol. Ges., Berlin,* 1878, No. 4-5 (Dec. 20, 1878). Reprinted in: Munk, H., *Über die Functionen der Grosshirnrinde.* Gesammelte Mittheilungen. Berlin: Hirschwald, 1880; 2nd edition, ibid., 1890. (Chapter IV, pp. 45-62 of 2nd edition.)

31. ———. Über die Functionen der Grosshirnrinde. Elfte Mittheilung: Stirnlappen von Hund und Affe. *Sitzber. Akad. Wiss., Berlin,* 1882. (July 20, issued July 27, 1882). Reprintd in: Munk, H., *Über die Functionen der Grosshirnrinde.* Gesammelte Mittheilungen, 2nd edition Berlin: Hirschwald, 1890. (Pp. 138-178.)

32. ———. Über die Ausdehnung der Sinnessphären in der Grosshirnrinde. Zweite Mitteilung. *Sitzber. Akad. Wiss., Berlin,* 1900, 770-793. Dritte Mitteilung. ibid., 1901, 1149-1183. Reprinted as Chapters X and XI, pp. 178-235, in Munk, H.: *Über die Functionen von Hirn und Rückenmark.* Gesammelte Mitteilungen. Neue Folge. Berlin: Hirschwald, 1909. Pp. 1-371.

33. MUSKENS, L. J. J. Das Supra-Vestibuläre System. Amsterdam, 1934. Pp. XVIII & 557.

34. SILBERPFENNIG, J. Contributions to the problem of eye movements: III. Disturbances of ocular movements with pseudohemianopsia in frontal lobe tumors. *Confinia Neurol.*, 1941, **4**, 1-13.

35. TEUBER, H.-L. Some observations on the organization of higher functions after penetrating brain injury in man. In: *The Biology of Mental Health and Disease.* New York: Hoeber-Harper, 1952.

36. TEUBER, H.-L., BATTERSBY, W. S., & BENDER, M. B. Changes in visual searching performance following cerebral lesions. *Amer. J. Physiol.*, 1949, **159**, 592.

37. ———. Performance of complex visual tasks after cerebral lesions. *J. Nerv. & Ment. Dis.*, 1951, **114**, 413-429.

38. TSCHERMAK-SEYSENEGG, A. v. Physiologisch-optische Studien. *Documenta Ophthalmologica*, 1948, **2**, 10-91. (See esp. 69-80.)

39. WERNER, H., & WAPNER, S. Experiments on sensory-tonic field theory of perception: IV. Effect of initial position of a rod on apparent verticality. *J. Exp. Psychol.*, 1952, **43**, 68-74.

40. WILCOXON, F. Some rapid approximate statistical procedures. New York: American Cyanamid Company, 1949. Pp. 1-16.

41. WITKIN, H. A. Perception of body position and of the position of the visual field. *Psychol. Monog.*, 1949, **63**, No. 7 (Whole number 302). Pp. III & 46.

42. WITKIN, H. A., & ASCH, S. E. Studies in space orientation: III. Perception of the upright in the absence of a visual field. *J. Exp. Psychol.*, 1948, **38**, 603-614.

*Department of Psychiatry and Neurology*
*New York University College of Medicine*
*477 First Avenue*
*New York City 16*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 177-190.

# PERFORMANCE ON A FORMBOARD-TASK AFTER PENETRATING BRAIN INJURY[*][†]

*Psychophysiological Laboratory, New York University-Bellevue Medical Center*

---

HANS-LUKAS TEUBER AND SIDNEY WEINSTEIN

---

## A. THE PROBLEM

No experiments are needed to convince us that man and apes take delight in fitting pegs into holes (30). The popularity of formboard tests, such as the one developed by Seguin (43, 44)[1] and modified by Goddard (13, 14) is therefore easy to understand. With a minimum of verbal instruction, most testees work away and produce scores which most testers are ready to accept as an index of "non-verbal intelligence" (2, 3, 8, 11, 39, 60). In order to obtain such an index, formboard tests have been given to children (1, 8, 49), including the very bad (7, 15, 16, 22, 29) or very dull (22, 29, 37), to adults literate (49, 58) and illiterate (61), to African tribes (36, 38) and other "natives" (12, 21, 25, 26, 32), and to aphasic (59) as well as non-aphasic patients who had brain tissue removed from various lobes (19, 23, 34, 59).

While the test is easily applied, it is far less easily interpreted. Superior performance may reflect high degrees of manipulative skill, of "visuo-spatial" ability, or of other capacities equally difficult to define, even with the help of factor analysis (56, 57). By the same token, performance on this test can be expected to be vulnerable to cerebral lesions in different areas of the brain, and for widely differing reasons. Breakdown in some hypothetical "spatial organization," difficulties with recall (of forms and positions explored), as well as more obviously "motor" deficits ("dyspraxia"), all may interfere with normal performance after loss of brain substance. Deficits on the test have to be accepted as a complex symptom of cerebral dysfunction; but it is meaningful for this very reason to apply the test after brain injury, in order to see whether this complex performance is equally disrupted by injury in any cerebral region, or whether there exist areas in the brain which are of maximum importance for success on this task.

---

[*] Received in the Editorial Office on April 1, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[†] The program of research of which this study is a part is supported by the Commonwealth Fund of New York, and by the Office of the Surgeon General (Department of the Army), Contract DA-49-007-MD-312.

[1] Seguin introduced his formboard "for the psycho-physiological training of the idiotic hand." Use of formboards as educational toys antedates their use as performance tests (13, 22, 37, 48) by half a century.

To be sure, regional localization of a symptom (such as a deficit on the formboard) is no substitute for an analysis of functions involved in the test (50); but regional localization, where possible, may be considered as a biological alternative to factor analysis (31, 51) suggesting what kinds of performance are impaired together, or preserved together, following a circumscribed lesion in a given region of the brain. The fact that the formboard test, as usually applied, depends only to a very small extent on overt language, makes it particularly valuable in this respect. The test can be applied to nearly every brain-injured *S* whether aphasic or not, and irrespective of visual field defects. The only possible restriction is severe bilateral motor deficit, a rather rare occurrence after traumatic, penetrating brain injury.

Our search for regional localization of formboard deficits was prompted, in particular, by Halstead's report (19) that damage of the frontal regions of the brain produces significantly greater impairment on the formboard than does comparable damage in any other area. In fact, Halstead has used scores from a modified Seguin-Goddard formboard procedure as an "indicator" of frontal lobe involvement. There is evidence, however, that men with right temporal lobe resections may show greater impairment on the formboard than men with frontal lobe resections (23, 34). Quite generally, one may question the notion expressed, independently, by Goldstein (17, 18), Rylander (41, 42), and Halstead (19), that the more complex a function, the more dependent it is on the integrity of the frontal lobes. A recent series of studies has cast doubt on this pre-eminence of the frontal lobes, both for monkey (6, 10, 20, 27, 28, 40) and man (24, 33, 34, 46, 52). There is increasing evidence that performance on many tasks involving complex sensory discrimination depends on temporal, posterior parietal, or parieto-temporo-occipital structures (4, 5, 6, 51, 52).

The present study was therefore undertaken in order to compare effects of penetrating injury (by missile) in various lobes of the brain, on performance on the Seguin-Goddard formboard. Specifically, we wished to test the importance of frontal, as compared with other lesions. Furthermore, we wanted to see whether different aspects of the task are differentially sensitive to focal brain injury.

## B. Subjects

The *S*s for this study were drawn from the 185 men with proven loss of brain substance (missile wounds) accessible to this laboratory, on whom our previous studies were based (51, 52). As in these previous studies, controls were chosen from 101 men with missile wounds of peripheral nerves, but without history of brain injury. All men (brain-injured and controls) had sustained their injuries in action 5-8 years prior to testing. For the present study, 35 men with brain injury and 12 controls (peripheral nerve injury) were tested. The 35 brain-injured included three men with aphasia of varying degrees. The control group (unilateral nerve injury cases) consisted of nine men with arm injuries and three with leg injuries.

The only known factor operating in selecting *S*s from the parent population was our concern in obtaining approximately equal numbers of *S*s with injuries in the anterior or posterior third of the brain, and in the right or left hemispheres. For purposes of analyzing results, *S*s were grouped in three ways.

### 1. *Longitudinal Subdivision*

All cases of brain injury were classified as having anterior (*a*), intermediate (*i*), or posterior (*p*) lesions. This grouping was accomplished by placing a transparent template on diagrams of the lateral aspect of the skull and brain; on the diagrams, the man's bone defect was marked, as were intracerebral foreign bodies (if any), and areas of known destruction of tissue. This information had been gathered previously for every case from roentgenographic evidence, and from surgeons' notes taken during debridement and cranioplasty. A grid on the template subdivided the major longitudinal axis of the brain into three equal sections, an anterior third (*a*), an intermediate third (*i*) and a posterior third (*p*). If any injury involved more than one section, the case was accordingly labeled as *ai* (anterior and intermediate thirds involved), *ip*, or *aip*. In grouping data, scores for cases with such mixed classifications were added into all subgroups they comprised. For example, in comparing *a, i,* and *p* groups with non-*a,* non-*i,* and non-*p* groups respectively, any *ai* score was added to the *a,* the *i,* as well as the non-*p* groups. The resulting classification of the 35 cases of brain injury yielded 17 *a,* 15 *i,* and 15 *p* cases.

### 2. *Classification by Lobes*

A second classification of cases was accomplished by means of another template, which marked off the major cerebral lobes, on the lateral aspect of the brain. Thus, we obtained cases with frontal (F), parietal (P), temporal (T), and occipital (O) injury, and various combinations of these. The resulting classification of the 35 cases, by major cerebral lobes involved, yielded 19 frontals, 13 parietals, 15 temporals, 11 occipitals, and one case of cerebellar injury.

### 3. *Classification by Laterality*

A third classification, supplementing the other two, was made according to the hemisphere involved: whether right (*r*) or left (*l*), unilateral (*u*) or bilateral (*b*). Table 1 gives a breakdown of the population according to locus of brain injury, whether in one or several lobes, and whether unilateral or bilateral.

TABLE 1
DISTRIBUTION OF PATIENTS BY LATERALITY AND LOCUS OF CEREBRAL LESION

| Unilateral lesions | | Bilateral lesions | |
|---|---|---|---|
| Right hemisphere | Left hemisphere | Right hemisphere | Left hemisphere |
| 3 F | 1 F | 6 F | F |
| 1 P | | | |
| 1 T | 2 T | | |
| 1 O | | 2 O | O |
| 1 FP | 2 FP | | |
| 1 FT | 1 FT | 1 FT | F |
| 1 TO | | 1 T | O |
| 1 PO | | 1 PO | O |
| 1 cbllm | | | |
| | 3 FPT | | |
| | 2 PTO | 1 PTO | O |
| | | 1 O | PTO |
| 11 Total | 11 Total | 13 Total | |

## C. PROCEDURE

A Seguin-Goddard formboard was placed on a table, resting on its longer edge, at a 60° angle. The 10 forms were numbered from 1-10, and placed within *E*'s reach on the table. *S* was seated, blindfolded, at the table; he was never permitted to see either board or forms. *E* handed the patient one form at a time, with the instruction to place it as rapidly as possible in the appropriate opening in the board. The first *S* was given the forms in order 1-10; in the subsequent presentations, the order was rotated, beginning with 2 and ending with 10, 1, for the next *S*, and so on. *S*s were permitted to use either hand or both. Four of the brain-injured men, and four of the controls, were limited to the use of one hand, the other being paralyzed by brain or nerve injury. *S*'s performance was closely observed, and every "false entry" (i.e., every attempt to fit a form into a wrong opening) was scored as an error. In addition to the error score, the time was noted from the beginning of each trial until the form was placed into the correct opening, or until five minutes (the maximum time score recorded) had elapsed. After each trial, the form was removed, and the next form handed to the patient.

After all 10 trials, *S* (still blindfolded) was allowed to rest for a brief period. The board was inverted (rotated 180°); *S* was told to resume his task, but he was not told that any change in the position of the board had occurred, nor was he informed of the fact that he had to work with the same 10 forms as before. This second run of 10 trials (II) was scored for time and errors, just as the first run (I). On completion of the second run, formboard and forms were taken out of sight; *S* was asked to remove his blindfold and to draw, on a blank sheet of paper, all those forms which he could remember. The number of pieces correctly remembered constituted the recall-score.

## D. Results

Results were analyzed in terms of the three primary scores, viz., time, error, and recall. Means and standard deviations were obtained for the various subgroups, *F* ratios computed, and *t* tests performed. Wherever significant *F* ratios were found, we applied Snedecor's modification of the *t* test (47), which corrects for inequality of variance. Significance between means was determined in all cases by means of two-tailed *t* tests. Transfer was

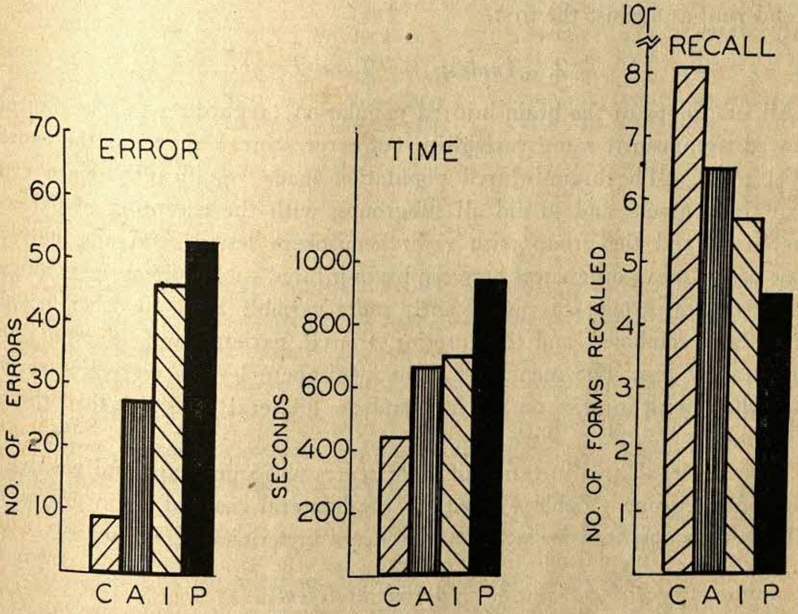


FIGURE 1
Error, Time, and Recall Scores for Controls (*C*), and for Subjects With Anterior (*A*), Intermediate (*I*), and Posterior (*P*) Brain Injury

assessed by comparing time or error scores for the first run (board upright) with corresponding scores for the second run (board rotated 180°). Results for the longitudinal subdivisions are given in Figure 1; those for the various lobes (omitting the one case of cerebellar injury) are presented in Tables 2-5.

### 1. *Analysis for Total Time*

The brain-injured group as a whole showed significantly greater variability of time scores than did the controls (Table 2). The same was true for all of the subgroups among the brain-injured, irrespective of their classification (whether *a, i, p; F, P, T, O; r, l, u,* or *b*). All of these sub-groups, with the exception of men with intermediate lesions (*i*-group), and men with lesions of the right hemisphere (*r*-group) took significantly more time than did the control group.

All groups tended to take less time on the second run as compared with the first, i.e., they showed positive transfer, in terms of time (see Table 4). The only exception was presented by the parietal group who showed slightly negative transfer. However, the transfer scores were significant in only one of these comparisons: the anterior group took significantly less time for the second run, as against the first.

### 2. *Analysis for Total Errors*

All subgroups of the brain-injured population (regardless of classification) showed significantly more variability (of error scores) than did the controls (Table 3). The brain-injured population made significantly more errors than the controls, and so did all subgroups, with the exception of parietals, temporals, and the group with right hemisphere lesions. Again, however, none of the mean differences between brain-injured subgroups were significant. The temporal group was significantly more variable than all other brain-injured cases combined, and the anterior, frontal, parietal, and occipital groups significantly less. The men with right hemisphere lesions were less variable than those with injuries on the left, and the bilateral cases less than the unilaterals.

Transfer evaluated in terms of error scores was significant and positive for the control group (Table 4) and for the bilateral cases of brain injury. In all other groups, transfer was not significant in terms of error scores.

### 3. *Analysis of Recall*

All brain-injured groups recalled significantly less forms than the controls (Table 5). None of the brain-injured groups differed from the control group

TABLE 2

COMPARISON AMONG GROUPS FOR TOTAL TIME IN SECONDS (POSITIONS I + II)

| Group | Mean | $N$ | $\sigma$ | | Group | Mean | $N$ | $\sigma$ | $t$ | $F$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Periph. | 442.2 | 12 | 112.2 | vs. | BI | 803.3 | 35 | 482.2 | 4.06**** | 18.5**** |
| | | | | | Front. | 736.6 | 19 | 443.4 | 2.83*** | 15.6**** |
| | | | | | Pariet. | 810.2 | 13 | 408.4 | 3.07*** | 13.2**** |
| | | | | | Temp. | 762.3 | 15 | 460.6 | 2.54** | 16.9**** |
| | | | | | Occip. | 876.2 | 11 | 510.7 | 2.84*** | 20.7**** |
| Front. | 736.6 | 19 | 443.4 | vs. | Non-Front. | 888.9 | 15 | 509.7 | .77 | 1.3 |
| Pariet. | 810.2 | 13 | 408.4 | vs. | Non-Pariet. | 799.8 | 21 | 518.0 | 1.55 | 1.6 |
| Temp. | 762.3 | 15 | 460.6 | vs. | Non-Temp. | 802.8 | 19 | 526.6 | .24 | 1.3 |
| Occip. | 876.2 | 11 | 510.7 | vs. | Non-Occip. | 769.2 | 23 | 461.0 | .64 | 1.2 |

*.02 $<P<$.05.
**.01 $<P<$.02.
***.001$<P<$.01.
**** $P<$.001.

TABLE 3

COMPARISON AMONG GROUPS FOR TOTAL ERRORS (POSITIONS I + II)

| Group | Mean | $N$ | $\sigma$ | | Group | Mean | $N$ | $\sigma$ | $t$ | $F$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Periph. | 8.8 | 12 | 4.7 | vs. | BI | 42.2 | 35 | 64.6 | 2.97*** | 189.9**** |
| | | | | | Front. | 31.7 | 19 | 35.1 | 2.90*** | 56.0**** |
| | | | | | Pariet. | 25.5 | 13 | 34.4 | 1.71 | 53.8**** |
| | | | | | Temp. | 41.1 | 15 | 72.2 | 1.68 | 237.0**** |
| | | | | | Occip. | 26.4 | 11 | 15.9 | 3.59*** | 11.5**** |
| Front. | 31.7 | 19 | 35.1 | vs. | Non-Front. | 41.5 | 15 | 70.6 | .48 | 4.0*** |
| Pariet | 25.5 | 13 | 34.4 | vs. | Non-Pariet. | 42.5 | 21 | 61.8 | 1.00 | 3.2** |
| Temp. | 41.1 | 15 | 72.2 | vs. | Non-Temp. | 32.0 | 19 | 32.6 | .45 | 4.9*** |
| Occip. | 26.4 | 11 | 15.9 | vs. | Non-Occip. | 40.6 | 23 | 63.5 | 1.04 | 16.0**** |

*.02 $<P<$.05.
**.01 $<P<$.02.
***.001$<P<$.01.
**** $P<$.001.

in variability of recall scores. However, some differences in variability appeared between brain-injured groups. The parietal and the posterior cases were significantly less variable than all others.

The mean recall scores for anterior cases and for frontals (see Table 5) were significantly better than for those without anterior or frontal lesions; by contrast, cases of posterior and occipital lesions were significantly worse in recall than those without such lesions.

## 4. *Interaction of Laterality and Lobes*

No significant interactions between laterality and lobe injured appeared in our data, with one important exception: cases of right temporal lesion were inferior to all other brain-injured groups with regard to transfer. Of the six men with right temporal lesions, five showed *negative* transfer (from

TABLE 4
SIGNIFICANCE OF TRANSFER SCORES FOR TIME (IN SECONDS) AND ERRORS

| Group | $M_{Diff.}$ | $N$ | $\sigma$ | $t$ |
|---|---|---|---|---|
| *Transfer Scores for Time* | | | | |
| Periph. | 33.0 | 12 | 112.2 | 1.34 |
| BI | 69.0 | 35 | 244.8 | 1.67 |
| Front. | 85.9 | 19 | 398.1 | 1.93 |
| Pariet. | — 9.1 | 13 | 428.1 | .14 |
| Temp. | 22.0 | 15 | 231.8 | .71 |
| Occip. | 81.3 | 11 | 615.7 | .78 |
| *Transfer Scores for Errors* | | | | |
| Periph. | 2.2 | 12 | 3.0 | 2.52** |
| BI | 5.1 | 35 | 29.9 | 1.00 |
| Front. | 1.6 | 19 | 31.4 | .46 |
| Pariet. | —2.4 | 13 | 26.6 | .60 |
| Temp. | —1.1 | 15 | 25.5 | .32 |
| Occip. | 2.5 | 11 | 19.8 | .74 |

*.02<P<.05.
**.01<P<.02.

TABLE 5
COMPARISON AMONG GROUPS FOR RECALL

| Group | Mean | $N$ | $\sigma$ | | Group | Mean | $N$ | $\sigma$ | $t$ | $F$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Periph. | 8.1 | 12 | 1.51 | vs. | BI | 5.5 | 35 | 2.2 | 3.94**** | .2 |
| | | | | | Front. | 6.3 | 19 | 2.1 | 3.00**** | .2 |
| | | | | | Pariet. | 4.9 | 13 | 1.9 | 4.57**** | .2 |
| | | | | | Temp. | 5.2 | 15 | 2.0 | 4.14**** | .2 |
| | | | | | Occip. | 4.2 | 11 | 2.2 | 4.88**** | .2 |
| Front. | 4.9 | 13 | 2.1 | vs. | Non-Front. | 4.5 | 15 | 1.6 | 3.00*** | 1.6 |
| Pariet. | 4.9 | 13 | 1.9 | vs. | Non-Pariet. | 5.9 | 21 | 4.8 | .83 | 6.4**** |
| Temp. | 5.2 | 15 | 2.0 | vs. | Non-Temp. | 5.8 | 19 | 2.1 | 1.00 | 1.1 |
| Occip. | 4.2 | 11 | 2.2 | vs. | Non-Occip. | 5.9 | 23 | 2.2 | 2.43** | 1.0 |

*.02 <P<.05.
**.01 <P<.02.
***.001<P<.01.
**** P<.001.

Conditions I to II), in terms of increasing time *and* error scores. By contrast, only three of the remaining 28 brain-injured men showed such negative transfer. The $X^2$ (with Yates' correction) is significant at the .01 level.

## E. DISCUSSION

Our results in 35 men with missile wounds in various regions of the brain confirm the notion that performance on a formboard task is readily disturbed by such injuries. The brain-injured *S*s take more time, make more errors, and recall fewer forms than the controls. Our results, however, do not support the earlier observation by Halstead (19) that frontal damage leads to greater impairment than other lesions. In terms of time or error scores, the brain-injured subgroups do not differ significantly one from the other. With regard to recall scores, men with lesions in the anterior third of the brain performed significantly better than those with injury to other parts of the brain, and men with lesions in the posterior third performed significantly worse. Grouped by major lobes injured, frontals recall significantly more forms, and occipitals significantly fewer than the remaining brain-injured population. If one considers further that our frontal group (in contrast to all other brain-injured) showed positive transfer for time (they worked significantly more rapidly on the second "run" as compared with the first), one must conclude that the claimed preeminence of the frontal region, for this complex task, has been denied by our results. The performance can be disturbed as we indicated, by injuries in the various locations, although the posterior third appears somewhat more important (at least for recall) than the anterior third of the brain.

These findings, while in contrast with expectations based on Halstead's reports, are nevertheless in line with a number of recent studies. Only minimal disturbances, in many tests, were found after bifrontal topectomies (33); our own reports, on men with missile wounds of the brain, have generally indicated that men with penetration of the posterior third of the brain did as poorly as, and sometimes worse than, men with anterior lesions, on a variety of complex visual tasks (51, 52). In experiments on subhuman primates, complex visual and tactile discriminations have been found to depend more on posterior cerebral regions (parieto-temporo-occipital) than on frontal areas of comparable size (6, 10, 45). However, neither the animal work, nor our recent studies have made it altogether clear whether the parieto-temporo-occipital systems act as a unit, or whether they are composed of separable foci, different in importance for various aspects of complex sensorimotor tasks (6).

We have found that brain injury in different lobes produced differential effects on certain aspects of the task, but not on other aspects. Men with occipital lesions had the lowest recall score, and men with temporal lesions (in the right, but not left hemisphere) showed negative transfer between first and second run. The poor showing of our occipital subgroup, on recall, is of particular interest, since the recall score is based on the only part of the test in which the subject is not blindfolded. Our patients with posterior penetration have been noted, previously, to do poorly on certain visual tasks of various degrees of complexity (4, 5, 51, 52, 53, 54). The observation of inferior performance on formboard-recall extends the generality of our previous findings.

The recall score, to be sure, reflects various undefined functions. Failure may be attributed to any one or several of the following factors: (*a*) inability to obtain form impressions on palpation, (*b*) inability to retain the impression, once obtained, (*c*) more specifically "visual" difficulties, or (*d*) inability to "translate" a haptic percept into a visual form (by drawing the form, on "recall"). While our complex task does not lend itself to an analysis of such factors, the possibility of selective impairment in transfer, from touch to vision, should be kept in mind. The existence of such impairment in "intersensory" transfer (from one modality to another), is suggested by recent studies from this laboratory, in which complex tactile and complex visual tests were compared for groups of men with brain injuries in various regions (46).

We should finally consider the particular difficulty of men with right temporal lesions, on the second "run" of our task. The high incidence of negative transfer (increase in time and error scores) when the board had been rotated, speaks for an inability, in these cases, of utilizing sensory information independently of spatial orientation of the "discriminanda." It is difficult to say whether this deficit is in line with those found in subhuman primates with bitemporal lesions (9, 27, 28). Patients with selective involvement of both temporal lobes are very rare, and there is no such case in our population (185 cases of cerebral gunshot wound), from which the 35 cases for the present study were drawn. It might seem paradoxical that selective impairment should result from unilateral temporal lesion on the right side, rather than on the left. All of our 35 *S*s had been right-handed before their injury, and their right hemisphere presumably represented the "lesser" lobe. Nevertheless, the finding of a distinct "right temporal" deficit is not without parallel.

Well before our study, Hebb noted some impairment of performance on a number of visuo-constructive tasks, in a patient with a large right temporal lobe resection (23). Subsequently, Milner compared groups of patients with cortical removals from various lobes (for relief of epilepsy) and demonstrated a maximal impairment on the formboard and related tasks, in a right temporal group (31). In our own laboratory, two recent observations are possibly related to our present findings: first, men with temporal lesions (either side) did poorly on a series of complex tactile tasks (46) involving generalization, sorting, conditional reaction and matching from sample; second, men with right hemisphere lesions were found to make abnormally large errors in judging the visual vertical under various conditions of body tilt (55).

It should be remembered, however, that occipital and temporal lesions in the present study were differentiated only by certain scores on the formboard task, and not by others. The one consistent trend in all the data was a slight, relative superiority of the frontal group. This discrepancy between our findings and those reported by Halstead (19) is not likely to be due to any differences in procedure. It is much more likely that the discrepancy reflects some real difference in the populations studied: missile wounds in our group, as against neoplasms, in Halstead's group. A tumor in the frontal lobes may disrupt brain function in a manner quite different from the effects of a penetrating lesion. Even following surgical removal of the tumor, some of these effects may persist, accounting for the contrast between the marked frontal lobe symptoms reported by Rylander and by Halstead, and the much less conspicuous deficits noted after psychosurgery (33), or missile wound in man (51), or after experimental removals in subhuman primates (10, 20, 40).

## F. Summary and Conclusions

Performance on a modified Seguin-Goddard formboard task has been claimed to be readily disrupted by brain injury, and most markedly by frontal lesions. To test these claims, a formboard test was given to 35 men with penetrating brain injury of various lobes, and to 12 controls with peripheral nerve lesions. The formboard was administered to the *blindfolded S,* first in its normal position, and then once more after 180° rotation. Performance was scored for total time and errors. At the end of the test, the *S* drew all the forms he could recall.

Performance of the brain-injured men was significantly inferior to that of the controls in terms of errors, time, and recall. Differences among brain-injured subgroups were limited to recall-scores. In terms of recall, men with anterior lesions and frontals did significantly better, and men with posterior lesions and occipitals significantly worse than all other brain-injured *S*s. Men with right temporal lesions showed "negative transfer" from first to second run of the test. Generally, recall scores on the formboard task yielded larger and more consistent differences between brain-injured and normals than did time or error scores. The results contradict the claim that performance on the formboard (including recall) should be more dependent on integrity of frontal, rather than posterior regions; for cases of missile wounds of the brain, performance is at least as vulnerable to parietal, temporal, and occipital penetrations.

## References

1. Amos, H. Ontario School Ability Examinations: A Performance Test Prepared More Especially for Use Among Children Who Are Deaf, Whose Native Tongue Is Other than English or Who for Any Reason Are Lacking in Language Facility. Toronto, Canada: Ryerson Press, 1936.
2. Arthur, G. A Point Scale of Performance Tests: Vol. I, Clinical Manual. New York: Commonwealth Fund, 1930.
3. ———. A Point Scale of Performance Tests: Vol. II, The Process of Standardization. New York: Commonwealth Fund, 1933.
4. Battersby, W. S. The regional gradient of critical flicker frequency after frontal or occipital injury. *J. Exp. Psychol.,* 1951, **42**, 59-68.
5. Battersby, W. S., Bender, M. B., & Teuber, H.-L. Effects of total light flux on critical flicker frequency after frontal lobe lesion. *J. Exp. Psychol.,* 1951, **42**, 135-142.
6. Blum, J. S. (Josephine Semmes), Chow, K. L., & Pribram, K. H. A behavioral analysis of the organization of the parieto-temporo-preoccipital cortex. *J. Comp. Neurol.,* 1950, **93**, 53-100.
7. Bronner, A. F., Healy, W., Lowe, G. M., & Shimberg, M. E. A Manual of Individual Mental Tests and Testing. Boston, Mass.: Little, Brown, 1927.
8. Carl, G. P. A new performance test for adults and older children: The Carl Hollow Square Scale. *J. of Psychol.,* 1939, **7**, 179-199.
9. Chow, K. L. Conditions influencing the recovery of visual discriminative habits in monkeys following temporal neocortical ablations. *J. Comp. & Physiol. Psychol.,* 1952, **45**, 430-437.
10. Chow, K. L., Blum, J. S. (Josephine Semmes), & Blum, R. A. Effects of combined destruction of frontal and posterior "associative areas" in monkeys. *J. Neurophysiol.,* 1951, **14**, 59-71.
11. Dearborn, W. F., Shaw, E. A., & Lincoln, E. A. A Series of Form Board and Performance Tests of Intelligence. Harvard Univ., Harvard Monogr. in Education, Series 1; Studies in Educational Psychology and Educational Measurement, No. 4. Cambridge, Mass.: Grad. School of Education, Harvard Univ., 1923. Pp. 64.
12. Garth, T. R., & Smith, O. D. The performance of full-blooded Indians on language and non-language intelligence tests. *J. Abn. & Soc. Psychol.,* 1937, **34**, 376-381.
13. Goddard, H. H. The form-board as a measure of intellectual development in children. *The Training School* (June) 1912, **9**, 1-4.
14. ———. The adaptation board as a measure of intelligence. *Training School Bull.,* 1915, **11**, 182-188.
15. ———. The Criminal Imbecile. An Analysis of Three Remarkable Murder Cases. New York: Macmillan, 1922.
16. ———. The Kallikak Family. A Study in the Heredity of Feeble-Mindedness. New York: Macmillan, 1925.
17. Goldstein, K. Die Lokalisation in der Grosshirnrinde. *Handb. d. norm. u. pathol. Physiol.* (ed. Bethe et al.), 1927, *Vol. X,* 600-842.
18. Goldstein, K., & Scheerer, M. Abstract and concrete behavior: an experimental study with special tests. *Psychol. Monog.,* 1941, **53**, 1-151.
19. Halstead, W. C. Brain and Intelligence. A Quantitative Study of the Frontal Lobes. Chicago: Univ. Chicago Press, 1947.
20. Harlow, H. F. The effect of large cortical lesions on learned behavior in monkeys. *Science,* 1950, **112**, 428.
21. Havighurst, R. J., & Hilkevitch, R. R. The intelligence of Indian children as measured by a performance scale. *J. Abn. & Soc. Psychol.,* 1944, **39**, 419-433.

22. Healy, W., & Fernald, G. M. Tests for Practical Mental Classification. Lancaster, Pa.: Review Publishing, 1911.

23. Hebb, D. O. Intelligence in man after large removals of cerebral tissue: defects following right temporal lobectomy. *J. Gen. Psychol.*, 1939, **21**, 437-446.

24. ———. Man's frontal lobes: a critical review. *Arch. Neurol. & Psychiat.*, 1945, **54**, 10-24.

25. Herrick, D. J. A comparison of Brahman and Panchama children in South India with each other and with American children by means of the Goddard form board. *J. Appl. Psychol.*, 1921, **5**, 253-260.

26. Klineberg, O. An experimental study of speed and other factors in "racial" differences. *Arch. of Psychol.*, 1928, **15**, No. 93, 109.

27. Klüver, H., & Bucy, P. C. Preliminary analysis of functions of the temporal lobes in monkeys. *Arch. Neurol. & Psychiat.*, 1939, **42**, 979-1000.

28. ———. An analysis of certain effects of bilateral temporal lobectomy in the rhesus monkey with special reference to "psychic blindness." *J. of Psychol.*, 1948, **5**, 33-54.

29. Knight, M. W. A comparative study of the performance of feeble-minded and juvenile delinquents on the Arthur Performance Scale and the Stanford-Binet Test of Intelligence. *J. Juvenile Res.*, 1934, **18**, 5-12.

30. Köhler, W. The Mentality of Apes. New York: Harcourt, Brace, 1925.

31. Lashley, K. S. Coalescence of neurology and psychology. *Proc. Amer. Philos. Soc.*, 1941, **84**, 461-470.

32. Mann, C. W. A test of general ability in Fiji. *J. Genet. Psychol.*, 1939, **54**, 435-454.

33. Mettler, F. A. (*Ed.*). Selective Partial Ablations of the Frontal Cortex: A Correlative Study of Its Effects on Human Psychotic Subjects. New York: Hoeber, 1949.

34. Milner, B. Intellectual effects of temporal-lobe damage in man. Ph.D. Thesis, McGill University, 1952.

35. Monakow, C. v. Die Lokalisation im Grosshirn und der Abbau der Funktionen durch kortikale Herde. Wiesbaden: Bergmann, 1914.

36. Nissen, H. W., Machover, S., & Kinder, E. F. A study of performance tests given to a group of native African Negro children. *Brit. J. Psychol.*, 1935, **25**, 308-355.

37. Norsworthy, N. Psychology of mentally deficient children. *Columbia Con. Philos. Psychol.*, 1906, 111.

38. Oliver, R. A. C. Mental Tests for Primitive Races. In: *The Testing of Intelligence* (Ed. by H. R. Hanley). London: Evans, 1937.

39. Pintner, R., & Paterson, D. G. A Scale of Performance Tests. New York: Appleton, 1925.

40. Riopelle, A. J., Harlow, H. F., Settlage, P. H., & Ades, H. W. Performance of normal and operated monkeys on visual learning tests. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 283-289.

41. Rylander, G. Personality Changes After Operations on the Frontal Lobes. London: Oxford, 1939.

42. Rylander, G. Mental changes after excision of cerebral tissue: A clinical study of 16 cases of resections in the parietal, temporal, and occipital lobes. *Acta Psychiat. et Neurol.*, 1943, Suppl. 20. Copenhagen: Munksgaard, 1943.

43. Seguin, E. Idiocy and Its Treatment by the Physiological Method. New York: Wood, 1866.

44. ———. The psycho-physiological training of the idiotic hand. *Arch. Med.*, New York, 1879, **2**, 149-156.

45. SEMMES, J. Cortical organization in somesthesis: Effects of lesions in posterior associative cortex on somatosensory function in Macaca mulatta. *Comp. Psychol. Monog.*, 1951, **105**, 219-249.

46. SEMMES, J., WEINSTEIN, S., GHENT, L., & TEUBER, H.-L. Performance on complex tactual tasks afer brain injury in man: Analyses by locus of lesion. *Amer. J. Psychol.*, 1954, **67**, 220-240.

47. SNEDECOR, G. W. Statistical Methods (4th Ed.). Ames, Iowa: Iowa State Press, 1946.

48. SYLVESTER, R. H. The form board test. *Psychol. Monog.*, 1913, **15**, No. 4 (Whole No. 65).

49. TERMAN, L. M. The Measurement of Intelligence: An Explanation of and a Complete Guide for the Use of the Stanford Revision of The Binet-Simon Intelligence Scale. Boston, Mass.: Houghton, Mifflin, 1916.

50. TEUBER, H.-L. Neuropsychology. Chapter 3 in *Recent Advances in Diagnostic Psychological Testing.* Springfield, Ill.: Thomas, 1950.

51. ———. Some observations on the organization of higher functions after penetrating brain injury in man. In *Biology of Mental Health and Disease.* New York: Hoeber-Harper, 1952.

52. TEUBER, H.-L., BATTERSBY, W. S., & BENDER, M. B. Performance of complex visual tasks after cerebral lesions. *J. Nerv. & Ment. Dis.*, 1951, **114**, 413-429.

53. TEUBER, H.-L., & BENDER, M. B. Critical flicker frequency in defective fields of vision. *Fed. Proc.*, 1948, **7**, 123-124.

54. ———. Alterations in pattern vision following cerebral lesions. *J. Gen. Psychol.*, 1949, **40**, 37-57.

55. TEUBER, H.-L., & MISHKIN, M. Judgment of visual and postural vertical after brain injury. *J. of Psychol.*, 1954, **38**, 161-175.

56. THURSTONE, L. L. A Factorial Study of Perception. Chicago: Univ. Chicago Press, 1944.

57. ———. Multiple-Factor Analysis. Chicago: Univ. Chicago Press, 1947.

58. WECHSLER, D. Measurement of Adult Intelligence (3rd Ed.). Baltimore: Williams & Wilkins, 1944.

59. WEISENBURG, T. H., & MCBRIDE, K. E. Aphasia: A Clinical and Psychological Study. New York: Commonwealth Fund, 1935.

60. WHIPPLE, G. M. Manual of Mental and Physical Tests (2 vols.). Baltimore: Warwick & York, 1914-15.

61. YERKES, R. M. (*Ed.*). Psychological Examining in the United States Army: Memoirs of the National Academy of Sciences, Vol. 15. Washington, D. C.: Government Printing Office, 1925.

*Department of Psychiatry and Neurology*
*New York University College of Medicine*
*477 First Avenue*
*New York City 16*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 191-202.

# TACTILE EXTINCTION IN PARIETAL LOBE NEOPLASM*†

*Psychophysiological Laboratory, New York University-Bellevue Medical Center*

ERICH G. KRUEGER, PLINY A. PRICE, AND HANS-LUKAS TEUBER

## A. INTRODUCTION

Although sensory disturbances after lesions of parietal lobe in man have been studied in considerable detail, there is little agreement regarding the nature of the deficits, or their dependence upon locus or extent of the parietal lesion. One source of difficulty is the considerable resilience of somatosensory functions following lesions of certain types. Simple cortical resections, as those performed for removal of epileptogenic foci (Evans, 1935), are rather unproductive of conspicuous or lasting tactile deficits. Similarly, data on experimental ablations in subhuman primates, though sparse, seem to indicate a marked resistance of sensory function to disorganization by uncomplicated cortical removals (Ruch and Fulton, 1935; Blum, 1951).

Such relative absence of deficits following cortical resections is surprising if one adduces the more massive symptoms reported in the clinical literature. Clinical descriptions are based primarily on cases of vascular or neoplastic disease. Apparently not only locus and mass, but type of cortical lesion play a rôle in determining the nature of the residual sensory performance. Considerations of this sort place important restrictions on the classical doctrine of localization which ascribes specific symptoms, or even functions, to given cortical areas.

In the somatosensory sphere, the difficulties of interpreting symptoms are magnified by the numerous ways in which abnormalities may manifest themselves. In addition to the deficits traditionally sought for—the deficits of passive motion and position sense; of two-point discrimination and point localization; of graphesthesia and stereognosis—there are the peculiarities of simple thresholds for touch observable following cortical lesions.

---

*Received in the Editorial Office on April 1, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

†From the Sections of Neurosurgery and Psychology, Veterans Administration Hospital, Bronx, New York (E.G.K. and P.A.P.); and the Psychophysiological Laboratory, Department of Psychiatry and Neurology, New York University-Bellevue Medical Center (H.L.T.). Work aided in part by the Commonwealth Fund of New York. Reviewed in the Central Office, USVA. The opinions expressed in this article are those of the authors' and do not necessarily reflect any official views of the US Veterans Administration.

As Head (1920) and others after him have pointed out (Weizsäcker, 1923; Stein, 1923 a, b; Stein and Weizsäcker, 1926; Holmes, 1927; Bender, 1946), there may be increased variability of thresholds. Moreover, multiple stimuli, successively or simultaneously applied to different body parts, may interact to produce abnormally raised thresholds or make it impossible for the patient to report any stimulus applied to an affected part.

Multiple simultaneous stimulation of presumably affected and normal body parts has been found particularly productive of sensory deficits in clinical situations (Bender, 1945). With such stimulation, patients frequently fail to notice a stimulus impinging on an impaired area, as long as some other area is intercurrently stimulated (extinction).

The first use of the method of multiple simultaneous stimulation appears to have been made in the animal laboratory (Loeb, 1884). Loeb, in experiments with dogs who had partial resection of an occipital or frontal lobe, demonstrated what he termed "hemiamblyopia" in this manner: Two pieces of meat were presented to the animal, one on each side of the center of the visual field. Operated dogs would turn first to the lure in the half-field ipsilateral to the lesion and then to the lure in the other half.

Soon afterwards, Oppenheim (1885), and Bruns (1886) extended this method to clinical neurologic examinations, employing either visual targets or tactile stimulation. Subsequently, others used the method, especially Poppelreuter (1917) who employed it systematically in acute cases of occipital lobe injury. However, the importance of the method has been brought into focus by Bender and his co-workers (Bender 1945; Bender and Teuber, 1946) during and after the second World War. Bender's work was soon confirmed by Nathan (1946), Reider (1946), and Critchley (1949), who underscored the frequent appearance of the phenomenon in various acute forms of parietal or occipital lobe involvement. They all stressed the clinical (localizing) value of the symptom, and suggested that systematic application of the method might also lead to a better theoretical understanding of the cerebral basis of somatosensory function.

In fact, the enhanced interaction of stimuli in the damaged substrate might be merely an exaggeration of the normal interaction of simultaneous stimuli in the visual or tactile sphere, and need not be specific for any particular parietal or extraparietal locus. Nevertheless, Critchley has recently suggested that involvement of the right parietal lobe might lead more frequently to extinction than lesions in other areas (Critchley, 1949).

At present, however, there are no data to show why extinction is conspicuous in some cases of somatosensory defect and absent in others. The symptom again seems to be more frequent with vascular and neoplastic lesions (Ziegler, 1952), although it has been reported after surgical resections of postcentral gyrus for relief of phantom limb pain (Kolb, 1950). In the absence of any identifiable focus, the suggestion has been made (Hebb, 1950) that the phenomenon is primarily attributable to unstable or "irritative" lesions. While Hebb's notion would not diminish the clinical value of extinction as a sign of morbid process, it would restrict physiologic inferences which could be drawn from the clinical observations. The case to be described in the following might serve to underscore many of these difficulties in the physiologic interpretation of sensory symptoms.

## B. Case Report

### 1. *Preoperative Status*

I.B., a 54-year-old, white, dry-goods salesman, was admitted on 8/9/51 to the Bronx VA Hospital; with two-weeks history of right-sided headaches, progressive weakness of lt. arm and lt. leg; and impairment of skilled movements in the lt. hand. On a previous admission from 4/24/51 to 5/17/51, a diagnosis of a rt. upper lobe lung lesion, apparently neoplastic in nature, had been made. The patient, who aside from mild diabetes had shown no other abnormalities, refused a chest operation. On re-admission the lung lesion was again demonstrated on roentgenograms.

*Neurological Examination* revealed a lt. lower facial weakness and a mild lt. hemiparesis, greater in ankle and wrist than in hip and shoulder. Finger-to-nose and heel-to-knee tests were poorly executed on the left. Deep reflexes were slightly increased on the left, and the lt. abdominal reflexes were absent. On 8/28/51 (three weeks after admission), a lt. Babinski and Chaddock sign could be demonstrated. At that time early bilateral papilledema with blurring of the nasal margins was noted. From the time of admission a left-sided homonymous but incongruent hemianopsia was found on visual field tests. Skull x-rays were negative. Psychometric examination, before operation, on the Wechsler-Bellevue Scale of Adult Intelligence gave the impression of subnormal functioning. The patient obtained a full scale *IQ* of 86, with a verbal quotient of 108 and a performance quotient of 67.

### 2. *Preoperative Sensory Deficits*

Repeated examination of the somatosensory system revealed no gross abnormality to cotton wisp and pin, if sensation was tested in the traditional fashion. However, the patient's judgments of passive movement, and of position passively assumed were slightly defective in the distal joints of the lt. hand, and more obviously defective in the lt. toes. Two-point discrimination was impaired in the lt. upper and lower extremities as compared to the right. There was slight impairment in recognition of figures traced on the skin over the lt. side; and some impairment, yet no complete loss, of ability to recognize common objects held by the lt. hand. Thus, a paper clip was hesitantly described as a "pin" when held by the lt. hand, but promptly identified as a paper clip when held by the right. A five-cent piece was called "some piece of metal—maybe a coin," on the left, and "a nickel" on the right.

All these sensory signs elicited on single stimulation (one part of the body at a time) were mild and could not always be demonstrated. By contrast, double simultaneous stimulation revealed a constant and conspicuous abnormality. Simultaneous contact of homologous parts by cotton wisps or pins led to "extinction" (i.e., failure to report) for any stimulus impinging on the patient's body to the left of the midline. This extinction involved even the lt. half of the patient's face. Asking the patient to "direct his attention" to the lt. side, or increasing the intensity of the stimulation on the left, were ineffective in overcoming extinction.

The definite sensorimotor signs, and particularly the persistent occurrence of extinction over the lt. side of the body, were held to be of sufficient localizing value for a diagnosis of a right parietal neoplasm, probably arising from a bronchogenic carcinoma.

### 3. *Craniotomy*

On 8/30/51, a rt. parieto-occipital craniotomy was performed under general endotracheal anesthesia. The posterior limb of the scalp incision was carried from the occipital protuberance upward along the sagittal line to a point 2cm. anterior to the interauricular line. The anterior limb extended from here downward in rectangular fashion to end 2cm. above and anterior to the rt. ear. On reflection of the dura, a surface tumor of firm granular consistency was seen in the area of the inferior parietal lobule. The inferior border of the tumor was at a distance of 5cm. above the tentorium and the posterior border at a distance of 3cm. from the posterior limb of the craniotomy. The circumscribed oval-shaped tumor measured 3x2.5 cm. on the surface; occupying the area of the angular gyrus and extending to the supramarginal gyrus. A brain needle advanced from a midpoint of the posterior limb of the exposure towards the mass struck the medial inferior surface of the tumor at a depth of 4cm. Here, 2 ccm. of yellowish cystic fluid, mixed with softened cerebral material, was aspirated. An oval cerebral block excision 5x4 cm. was carried out; and a nodular tumor extending to a depth of 3cm., and of approximately golf ball shape, was enucleated. The weight of the removed specimen was 29 Gms. The histological diagnosis was: Metastatic Carcinoma. In trimming the edges of the cerebral wound, some additional cerebral tissue was removed with the sucker. Following hemostasis, the dura was closed, the bone-muscle flap was replaced and the wound closed in layers. The patient received 1000 cc. of blood during the procedure.

### 4. *Postoperative Course* (*Early Phase—Improved Sensory Status*)

On the morning of the 1st postoperative day the patient was alert, speaking freely and moving all extremities; except for the mild weakness of lt. arm and leg noticed before operation. Extinction on double stimulation by cotton wisp was still present over the entire body (including face) to the left of the midline, but response to double stimulation by pins had changed. Extinction was no longer found over the trunk, in the face, in the proximal third of the lt. upper extremities, or in the proximal part (above the knee) of the lt. leg.

In the afternoon of the same day, somewhat more detailed sensory tests were performed revealing further improvement in sensory status. Extinction to light touch (cotton) was absent in the face and inconstant over the trunk and in the lower extremity. However, extinction was still present on every trial in the distal third of the lt. upper extremity.

These residual sensory changes could also be demonstrated by noting "disappearance times" (adaptation) for light touch or pain (Table 1).

This abnormally short "tactile adaptation" was associated with dysgraphesthesia in the lt. hand (Table 2).

Recognition of solid objects by the lt. hand was still somewhat defective; for example: Teaspoon in lt. hand: "This is a piece of metal, a smooth object; could be a shoehorn." Teaspoon in rt. hand: "This feels like a spoon."

There was mild disorientation for place and time. Fundi and visual fields were unchanged and remained so throughout the entire postoperative course.

TABLE 1

| | Latency | | Adaptation | |
|---|---|---|---|---|
| | Lt. | Rt. | Lt. | Rt. |
| Pin on volar distal surface of third finger: average of three trials: | 2 sec. | 0 sec. | 7½ sec. | 80 sec. |
| Pin on volar distal surface of first toe: | 1 sec. | 1 sec. | 8 sec. | 81 sec. |

TABLE 2

FIGURE WRITING (PALMS)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Left* | | | | | | | | |
| Figures given: | 3 | 3 | 5 | 5 | 4 | 5 | 3 | 4 |
| Response: | 6 | ? | 4 | 4 or 7 | 2? | ? | 8 | 5 |
| *Right* | | | | | | | | |
| Figures given: | 3 | 4 | 5 | 3 | 4 | 5 | 3 | 4 |
| Response: | 2/3 | 4 | 5 | 5? | 4 | 5 | 3 | 4 |

### 5. *Further Postoperative Course: Disappearance of Sensory Signs*

On the 2nd and 3rd postoperative days the sensory status remained unchanged. On the 4th day, however, extinction could no longer be elicited on double stimulation over the entire body, irrespective of type of stimulus (cotton or pin). Adaptation times for continued contact and pinprick had likewise become normal (longer than 60 sec.) on lt. fingers and toes, and were equal to those obtained for the corresponding parts of the patient's rt. side. Ability to judge passive motion and position of distal portions of limbs had also become normal and equal on either side. The sensory status continued to be normal; although the mild hemiparesis, the incomplete hemianopsia, and moderate disorientation persisted without change. Psychometric examination (6 weeks after operation) on the Wechsler-Bellevue Scale (Form II) yielded a verbal *IQ* of 116, a performance *IQ* of 83, and a full-scale *IQ* of 98.

### 6. *Attempts at Re-establishing the Sensory Symptoms Under Sodium Amytal*

On the 12th postoperative day (9/11/51), it was decided to inject the patient with sodium amytal. Earlier reports on other cases have indicated that a hemisensory syndrome which has disappeared in the course of restitution can be reinstated by the injection of barbiturates (Bender, Shapiro, and Teuber, 1949). Contrary to these earlier reports, injection of sodium amytal (.3 Gm. in 3cc. water, i.v.) failed to produce any sensory dysfunction; although the characteristic drug effects of nystagmus on forward gaze, slurring of speech, and mild euphoria were noted. Immediately following injection, and for 30 min. afterwards, tests of sensory functions were carried out including evaluation of response to light touch and pin, on single and double stimulation; evaluation of sense for position and passive motion; figure writing on both palms; and two-point thresholds. Transverse two-point thresholds were obtained (Table 3).

TABLE 3

| | *Left* | *Right* |
|---|---|---|
| Palm of hand: | 2.5 cm. | 2.5 cm. |
| Sole of foot: | 2.9 cm. | 3.5 cm. |

A second injection of sodium amytal three weeks after operation had similarly negative results, despite an increase in dosage to .6 Gm. (in 6cc. water).

### 7. *Terminal Course*

Four weeks after operation the patient began to complain of weakness in the rt. arm and leg, and of intermittent pain referred to the deep tissues of the rt. thigh. A pneumoencephalogram was performed revealing slight symmetrical dilation of ventricles without any shift. On 10/22/51, a rt. upper lobectomy for removal of the primary pulmonary lesion was carried out. Recovery from this operation was uneventful, but the motor symptoms continued to grow more obvious. The patient became increasingly disoriented, and developed a marked dysphasia. At the end of his hospital stay, he had a complete rt. hemiplegia, was mute, and incontinent of urine. He expired on 1/5/52, about four months after craniotomy, of a bilateral bronchial pneumonia.

### 8. *Post Mortem Findings*

The autopsy showed no recurrence of metastases in the rt. parieto-occipital area; there were, however, two smaller metastatic tumors of cherry size in the rt. premotor area, and one in the lt. prefrontal area; as well as a larger egg-sized tumor in the lt. parasagittal fronto-parietal region. There were also multiple metastases in the cerebellar hemispheres.[1] Examination of the lungs revealed that the bronchogenic carcinoma had not radically been removed.

## C. COMMENT

This case illustrates the difficulties of establishing one-to-one relations between an existing cortical lesion and seemingly corresponding changes in somatosensory function. As long as the space occupying lesion was present in the posterior parietal lobe, the patient exhibited classical signs of sensory impairment, viz., diminished position sense, dystereognosis, and diminished discrimination of two compass points simultaneously applied.

Shortly upon removal of the tumor (an operation which left a cortical lesion 5x4 cm. in the rt. posterior inferior parietal lobe), all these symptoms disappeared. It is noteworthy that before the removal the patient exhibited extinction on bilateral tactile stimulation. Extinction of sensations on the lt. side, upon concomitant stimulation of the right, was so marked before operation that it had constituted the major localizing sign suggesting a rt. parietal lesion. Nevertheless, extinction disappeared pari passu with the other sensory signs after surgery.

---

[1] Histological reports will be presented in a separate publication.

Contrary to expectations based on other cases, neither extinction nor any of the other sensory symptoms could be reestablished postoperatively under sodium amytal. The removal of pathological tissue, despite the unavoidable destruction of surrounding healthy cortex, had thus radically removed all abnormality of sensation. In view of such paradoxical improvement, it becomes hazardous to ascribe the extinction phenomenon and the associated sensory dysfunction, generally, to mere absence of tissue; the symptoms in this case at least, seemed clearly related to the presence of diseased tissue prior to operation.

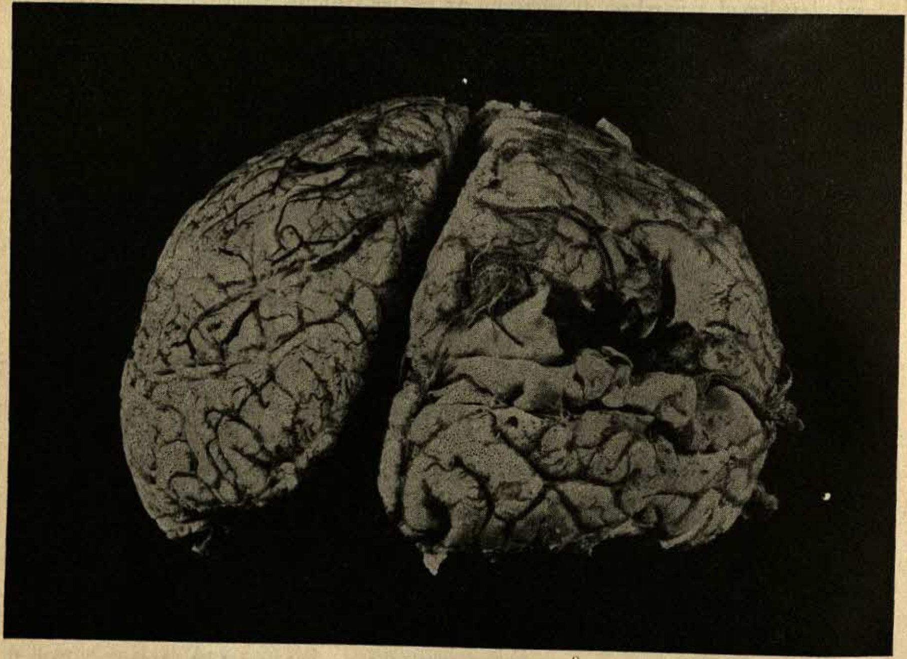


FIGURE 1
POSTERIOR VIEW OF AUTOPSY SPECIMEN SHOWING SHRUNKEN OPERATIVE DEFECT IN RIGHT INFERIOR PARIETAL LOBULE

It should be stressed, furthermore, that our case after operation presented a lesion in precisely that area to which Head attributed mediation of certain "higher" aspects of tactile sensitivity. According to Head (1920) inconstancy of touch thresholds, particularly on multiple (repeated or simultaneous) contact, depend uniquely on the invasion of supramarginal and angular gyri. Similarly, Evans (1935) in his review of 17 cases from Penfield's clinic attempted to show that lasting deficits of tactile discrimination (particularly for two compass points) was regularly found in men with cortical removals from the posterior parietal region. It would seem from Figure 1 that the operation had involved the very area implicated by Evans in complex tactile functions; yet we were unable to demonstrate any sensory abnormality in this patient following the operation despite the presence of massive sensory disturbances prior to surgery.[2]

Cases of paradoxical improvement of function in the injured brain, upon additional removal of cerebral tissue, have occasionally been reported. Analysis of Evans' own case reports (1935) suggests that his Case 13 (massive ablation from rt. temporal lobe) might have been similar in this respect. In this case two operations were performed. Three years after the first operation, impairment of weight discrimination in the patient's lt. hand could still be noted. Four years after the first operation, the patient was reoperated and the resection of the rt. temporal lobe was carried still further upward and backward than before. Ten months after the second operation weight discrimination was tested and found to be normal (Evans, p. 360).

In the motor system, similarly paradoxical improvement on additional removal of cortex has been noted by Welch and Penfield (1950) in two cases of spastic paralysis due to early trauma. Unpublished data by Semmes and Chow (in press) on seriatim ablations of motor cortex in macaques suggest that the phenomenon is not as rare as might be assumed.

In the sphere of general behavior and intellectual functioning, the classical instance of paradoxical improvement is the case recorded by Hebb and Penfield (1940). In their patient, removal of an epileptogenic bifrontal scar, together with apparently viable frontal tissue, converted a man with striking behavioral disturbances (childishness, obscenity, intellectual impairment), into one whose behavior had become unobtrusive, and whose tested intelligence was repeatedly found to be within normal limits.

---

[2] Our conclusion agrees with that reached recently by Penfield and Rasmussen (1950) who state: "Cases reported by Evans (1935) suggested that lesions in this area (posterior parietal) were capable of producing marked sensory deficits without encroachment upon the postcentral gyrus itself; but it seems evident now that if excision is carefully restricted to the grey matter by means of suction, no cortical sensory loss may be produced unless there is interference with the arterial supply or the subcortical connections of the postcentral gyrus."

The importance of such cases, in conjunction with the one here presented, lies in their didactic value. If more radical removal of tissue leads to improvement or disappearance of symptoms, then these symptoms cannot be unequivocally related to loss of tissue. The nature of the disorganizing influence exerted by a metastatic brain tumor is of course not understood, but the marked edema surrounding such lesions might play a rôle.

The observations do not detract from the pragmatic value of classical localizing signs, especially in the diagnosis of suspected space-occupying lesions. In the present case these signs, reinforced by the conspicuous appearance of extinction phenomena, permitted adequate preoperative diagnosis and led to the removal of the cerebral tumor.

The importance of the case lies in the fact that the area occupied by the tumor cannot be crucial for "higher" tactile function, despite the assumptions of Head and Evans. In addition, the case underscores that we know just as little about the physiological basis for tactile extinction as we know about that of the more traditional signs of sensory dysfunction in cerebral lesions. We still cannot explain why extinction appears when it does and why it fails to appear when it does not.

## D. SUMMARY

Description of a patient who presented disturbances of passive motion and position sense, tactile adaptation, two-point discrimination, graphesthesia and stereognosis; together with marked "extinction" of sensation on bilateral simultaneous stimulation. All these symptoms were found in the presence of a bronchogenic (metastatic) tumor in the rt. posterior inferior parietal region. Following radical removal of the tumor, all of the sensory symptoms, including the extinction phenomena, disappeared and could not be reestablished on administration of sodium amytal. The case serves to emphasize the reversible nature of "extinction" under certain conditions, and underscores the possibility that the presence of diseased tissue, as pointed out by Hebb (1950), may be more productive of symptoms than the mere absence of cerebral tissue.

## REFERENCES

1. BENDER, M. B. Extinction and precipitation of cutaneous sensations. *Arch. Neurol. & Psychiat.*, 1945, **54**, 1-9.
2. ———. Changes in sensory adaptation time and after-sensation with lesions of parietal lobe. *Arch. Neurol. & Psychiat.*, 1946, **55**, 299-319.
3. BENDER, M. B., SHAPIRO, M. F., & TEUBER, H. L. Allesthesia and disturbance of the body scheme. *Arch. Neurol. & Psychiat.*, 1949, **62**, 222-231.
4. BENDER, M. B., & TEUBER, H. L. Phenomena of fluctuation, extinction, and completion in visual perception. *Arch. Neurol. & Psychiat.*, 1946, **55**, 627-658.
5. BLUM, J. S. Cortical organization in somesthesis: Effects of lesions in posterior associative cortex on somatosensory function in Macaca mulatta. *Comp. Psychol. Monog.*, 1951, **20**, 219-249.
6. BRUNS, L. Ein Beitrag zur einseitigen Wahrnehmung doppelseitiger Reize bei Herden einer Grosshirnhemisphäre. *Neurol. Centralbl.*, 1886, **5**, 198-199.
7. CRITCHLEY, M. Phenomenon of tactile inattention with special reference to parietal lesions. *Brain*, 1949, **72**, 538-561.
8. EVANS, J. P. A study of the sensory defects resulting from excision of cerebral substance in humans. *Res. Publ. Ass. Nerv. & Ment. Dis.*, 1935, **15**, 331-366.
9. HEAD, H. Studies in Neurology (2 vols.). London: Oxford Med. Publications, 1920. Pp. ix & 862.
10. HEBB, D. O. Animal and physiological psychology. *Ann. Rev. Psychol.*, 1950, **1**, 173-188.
11. HEBB, D. O., & PENFIELD, W. Human behavior after extensive bilateral removals from the frontal lobes. *Arch. Neurol. & Psychiat.*, 1940, **44**, 421-438.
12. HOLMES, G. Disorders of sensation produced by cortical lesions. *Brain*, 1927, **50**, 413-427.
13. KOLB, L. C. Observations on the somatic sensory extinction phenomenon and the body schema after unilateral resection of the posterior central gyrus. *Trans. Amer. Neurol. Assoc.*, 1950, 138-140.
14. LOEB, J. Die Sehstörungen nach Verletzung der Grosshirnrinde. *Pflüger's Arch.*, 1884, **34**, 67-172.
15. NATHAN, P. W. On simultaneous bilateral stimulation of the body in a lesion of the parietal lobe. *Brain*, 1946, **69**, 325-334.
16. OPPENHEIM, H. Ueber eine durch eine klinisch bisher nicht verwertete Untersuchungsmethode ermittelte Form der Sensibilitätsstörung bei einseitigen Erkrankungen des Grosshirns (Kurze Mittheilung). *Neurol. Centralbl.*, 1885, **23**, 529-533.
17. PENFIELD, W., & RASMUSSEN, T. The Cerebral Cortex of Man. A Clinical Study of Localization of Function. New York: Macmillan, 1950. Pp. xv & 248.
18. POPPELREUTER, W. Die Psychischen Schädigungen durch Kopfschuss im Kriege, 1914-16. Leipzig: Voss, 1917. Pp. xvi & 210.
19. REIDER, N. Phenomena of sensory suppression. *Arch Neurol. & Psychiat.*, 1946, **55**, 583-590.
20. RUCH, T. C., & FULTON, J. F. Cortical localization of somatic sensibility: The effect of precentral, postcentral and posterior parietal lesions upon the performance of monkeys trained to discriminate weights. *Res. Publ. Ass. Nerv. & Ment. Dis.*, 1935, **15**, 289-330.
21. SEMMES, J., & CHOW, K. L. Motor effects of Rolandic and extra-Rolandic cortical ablations. (In preparation.)

22. STEIN, H. Die Labilität der Drucksinnschwelle bei Sensibilitätsstörung. *Dtsch. Z. Nervenheilk.*, 1923, **80**, 57-74.

23. ———. Nachempfindungen bei Sensibilitätsstörungen als Folge gestörter Umstimmung (Adaptation). *Dtsch. Z. Nervenheilk.*, 1923, **80**, 218-237.

24. STEIN, H., & WEIZSÄCKER, V. Ueber klinische Sensibilitätsprüfung. *Dtsch. Arch. klin. Med.*, 1926, **151**, 230-253.

25. WEIZSÄCKER, V. Ueber den Funktionswandel, besonders des Drucksinnes, bei organisch Nervenkranken und über Beziehungen zur Ataxie. *Pflüger's Arch.* 1923, **201**, 317-332.

26. WELCH, K., & PENFIELD, W. Paradoxical improvement in hemiplegia following cortical excision. *J. Neurosurg.*, 1950, **7**, 414-420.

27. ZIEGLER, D. K. Extinction phenomenon in patients with verified cerebral tumors. *Neurology,* 1952, **2**, 501-508.

*Department of Psychiatry and Neurology*
*New York University College of Medicine*
*477 First Avenue*
*New York City 16*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 203-209.

# INFANT SPEECH AS A POSSIBLE PREDICTOR OF LATER INTELLIGENCE*

*Department of Psychology, Fordham University*

FRANK L. CATALANO AND DOROTHEA MCCARTHY

## A. INTRODUCTION

The present study is a pilot investigation designed to explore the feasibility of using certain measures of infant speech in the prediction of later intellectual status. In the construction of intelligence tests, items sampling the verbal aspects of mental ability have proved by far the most valuable type of item. Attempts to measure intelligence before the age of 18 months to two years have been extremely disappointing from the standpoint of agreement with later tests of the same individuals (1, 2, 11). However, the age at which it is possible to secure some reasonably stable predictive measures of intellectual capacity is in the latter half of the second year of life (2, 3, 4). This is the period in which the most striking developments in the onset of true language occur in most normal children. The various infant tests however, have been heavily weighted with motor and sensori-motor items (3, 8, 22), and few items below one year of age represent a sampling of verbal or abstract ability which has proven such a valuable source of intelligence test items at higher age levels (24). At these upper ages measures of motor behavior do not correlate highly with measures of general intelligence. It should not be surprising therefore, that infant tests heavily weighted with readily observable motor behavior have not shown satisfactory validity.

The techniques developed by Irwin and his associates (6, 17, 18, 19, 20, 21) in their studies of infant speech however, suggest the possibility of a new approach to infant intelligence testing based on more refined observations of the verbal factor as it is found in the pre-linguistic babblings of infants. A previous study by Fisichelli (10)[1] under the direction of the senior author employed phonemic analysis of tape recordings of 30-minute samples of the utterances of 100 institutionalized infants. The subjects, who were originally examined in two foundling homes, ranged in age from six months to 18

---

*Received in the Editorial Office on April 15, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors are grateful to Dr. Fisichelli for making possible the identification of cases for purposes of this research and for permission to use the data supplied in her unpublished dissertation.

months at the time the recordings were made. There were 20 children at each of five discrete age levels, 6, 9, 12, 15, and 18 months. The data revealed a number of striking developmental trends with age. For the most part the results tended to confirm the work of Irwin and his associates and indicated a marked retardation of institutionalized infants compared with those who were living in normal family situations studied by the Iowa investigators (6). The most promising measures were the frequency of consonant sounds (CF), the number of different kinds of consonant sounds or consonant types (CT), the consonant-vowel frequency ratio (CF/VF) and the consonant-vowel type ratio (CT/VT).

## B. Method

The present investigation is a follow-up study of 23 of the children studied by Fisichelli (10). One of the institutions which supplied two-thirds of the cases for the original investigation would not coöperate in the follow-up study, stating that it was against their policy to divulge the whereabouts of children no longer in the institution. Of the 35 cases originally contacted at the Angel Guardian Home[2] in Brooklyn 23 were available for retesting approximately three years later. Five of these were still in the institution and the others were in boarding homes. Eight cases which had either returned to their own families or had been legally adopted were not available for retest. Form *L* of the Stanford-Binet intelligence test was administered to all subjects in their boarding homes by the junior author. Those children who had not been placed in boarding homes were tested in the institution. Preliminary arrangements for testing the children in boarding homes were made with the boarding mothers by the social worker assigned to each case. Appointments were arranged for hours when other children living in the homes were at school or taking naps so as to insure a minimum of distraction.

All children were told of the examiner's visit in advance by their social workers. Coöperation was obtained from the boarding mothers, and in all instances the examiner found the children at home and ready to be examined at the time scheduled. In most cases the children were tested alone without the presence of the boarding mother. The examiner considered that adequate rapport and coöperation were obtained in all cases so that all examinations yielded valid results.

In Table 1 is shown the age distribution of the 23 cases available for intelligence tests in the follow-up study at the time of Fisichelli's (10) initial

---

[2] The writers are deeply grateful to Sister Leócadia of the Angel Guardian Home, Brooklyn, N. Y., for her gracious coöperation.

TABLE 1
AGE DISTRIBUTION AT TIME OF RECORDINGS OF SPEECH

| Age in months | 6 | 9 | 12 | 15 | 18 |
|---|---|---|---|---|---|
| Number of cases | 2 | 6 | 4 | 1 | 10 |

TABLE 2
AGE DISTRIBUTION AT TIME OF INTELLIGENCE TEST

| Age in months | 36-41 | 42-47 | 48-54 |
|---|---|---|---|
| Number of cases | 2 | 11 | 10 |

recordings of their pre-linguistic utterances. The mean age of this group at the initial observation was 13.3 months. By the time of the follow-up study an average interval of 31.3 months had elapsed since the infant language sample so that the average age was 44.8 months. Table 2 shows the age distribution for the same cases at the time the intelligence tests were administered. All children in the sample were talking intelligibly by the time the intelligence tests were given.

## C. RESULTS

Table 3 shows the distribution of *IQ*'s on the tests given; the range was from 57 to 102 with an average at 74.7. Only three cases earned *IQ*'s above 90 and four were below 60. It thus appears that this sample of children who spent much of their infancy in an institution and who were not released for adoption, but who were spending their preschool years in boarding homes were seriously retarded mentally as measured by a standard verbal test. It cannot be determined from the data available whether this retardation was due to poor native endowment or to institutional experience. Neither can it be determined whether these children were originally of normal potential or whether selective factors such as slowness in their early development contributed to their not being placed for adoption.

TABLE 3
DISTRIBUTION OF *IQ*'S ON STANFORD-BINET FORM *L*

| *IQ* range | 55-64 | 65-74 | 75-84 | 85-94 | 95-104 |
|---|---|---|---|---|---|
| Number of cases | 7 | 5 | 6 | 3 | 2 |

Using the results of Fisichelli's (10) recordings of infant vocalizations, which had been analyzed phonemically, correlations were obtained between *IQ*'s on the Stanford-Binet Form *L* at the later ages and various measures of infant speech shown in Table 4. All of the correlations obtained between

TABLE 4
CORRELATIONS BETWEEN STANFORD-BINET *IQ* AND VARIOUS MEASURES OF INFANT SPEECH

| Speech measures | Zero-order $r$'s* | Partial $r$'s, age at speech test constant | Partial $r$'s, age at both tests constant |
|---|---|---|---|
| Consonant types | .45 | .45 | .41 |
| CF/VF ratio | .38 | .41 | .41 |
| CT/VT ratio | .35 | .34 | .29 |
| Vowel types | .29 | .28 | .31 |
| Consonant frequency | .29 | .27 | .26 |
| Vowel frequency | .25 | .24 | .24 |
| Frequency of dental-alveolar consonants | .18 | .16 | .15 |
| Frequency of back vowels | .04 | .02 | .07 |

*For $N = 23$, a zero order $r$ of .413 is significant at the five per cent level of confidence, and one of .352 is significant at the 10 per cent level.

the various indices of infant speech and later *IQ* are positive. Measures of consonant development which in Fisichelli's data showed the most marked trends with age yield the highest correlations with later *IQ* (see Table 4). The correlation between *IQ* and the number of different kinds of consonants (CT) in the recordings of infant speech yields a raw coefficient of .45 which remains the same when age at first test is partialled out. When further adjustment is made for variation in age at the time of the intelligence testing (necessary because the interval between examinations was not constant for each individual) the correlation drops to .41.

When frequency of the various sounds is considered, regardless of the variety, positive but somewhat lower correlations are found for both vowels and consonants. When the ratio of consonant to vowel frequency is used however, a correlation of .41 with later *IQ* is again obtained, when age at both testings is held constant. This correlation is significant at the 10 per cent level of confidence ($t = 1.96$, midway between $P = .05$ and $P = .10$).

As shown in Table 1, 10 of the cases were first examined by Fisichelli for speech records at 18 months of age so that for them age at first testing was held constant experimentally rather than statistically. For this group the zero-order correlation between *IQ* and consonant types was .40 and that between *IQ* and consonant-vowel frequency ratio was .48.

Table 5 presents several multiple correlations between *IQ* and combinations of several of the more promising measures of infant speech namely, consonant types, consonant-vowel frequency ratio, and vowel types. As may be seen these values between *IQ* and various pairs of speech variables range from .46 to .48. When *IQ* was correlated with all three of the above speech

TABLE 5
MULTIPLE CORRELATIONS BETWEEN *IQ* AND CERTAIN COMBINATIONS OF MEASURES OF INFANT SPEECH

| $R_{IQ.CT\ CF/VF}$ | $R_{IQ.CT\ VT}$ | $R_{IQ.CF/VF\ VT}$ | $R_{IQ.CT\ CF/VF\ VT}$ |
|---|---|---|---|
| .46 | .48 | .47 | .52 |

variables the correlation was raised to .52. Due to the small number of cases in this pilot study there is appreciable shrinkage in the multiple correlations. However, the fact that the correlations remain in the neighborhood of .40, even when corrected for shrinkage, indicates that these measures of infant speech are indeed promising and deserving of further research on larger numbers of subjects.

## D. DISCUSSION

In interpreting data of this sort it must be remembered that the number of cases involved is too small to yield anything but suggestive results. It should also be noted that these correlations are undoubtedly lower than would be obtained on a group representative of the entire ability range. The present group was restricted in range of ability to the lower end of the distribution (*IQ* 57 to 102). If similar relationships are found to hold when data are available on a larger sample representing the entire range of ability the use of measures of infant speech for the prediction of later intelligence would be promising indeed.

The correlations obtained here are higher than most the writers have seen in the literature between infant intelligence tests and later *IQ*, even over much shorter intervals. Irwin and Chen (21), using an index of discrepancy between infant and adult speech sounds, found a correlation with *MA* of .40 for children one to six months, and one of .56 for children seven to 12 months of age. The correlations with *IQ* were however negligible or negative indicating that the measure employed was most closely related to *CA*. In their 19-24-month age group higher correlations of .46 and .42 were obtained with *IQ*. These values were found when the speech sample and the intelligence test were taken at the same age. It should be remembered however, that they employed measures of discrepancy or deviation of infant from adult speech sounds instead of positive measures of linguistic accomplishment in infancy. The Kuhlmann scale which they employed at the infant level is heavily weighted with sensori-motor items and has not proven a good predictor of intelligence at higher age levels.

In view of the facts that phonemic measures of infant speech have been found by Brodbeck and Irwin (6) to yield socio-economic group differences, even in the first six months of life which correspond to the differences in intellectual status of such groups at higher ages, that such measures differentiate between institutionalized children and those living in their own homes, and that they yield correlations of the magnitude reported in this study even with small groups restricted in range of ability, it would appear that they are indeed promising and deserving of further study as possible predictors of later intelligence. Of interest in this connection is the statement by Anderson (2) based on his careful study of 91 infants at the Brush Foundation. He stated in conclusion: "Of great importance are the data concerning the significant rôle that test items, involving the use of language and the understanding of words, play in the prediction of intelligence. For example, the correlation of a test composed of language items only is more significant than other classifications of items."

It appears then, that this type of investigation needs to be undertaken on a sufficiently large scale with a representative group of children in order to yield definitive results. If a large-scale repetition of the present study confirms our findings, it appears that infant intelligence tests of the future should incorporate measures of the verbal factor even at the earliest pre-linguistic stages. It is conceivable that measures of consonant types, and consonant-vowel frequency ratio in the babblings of young infants might be more valuable predictors of later intelligence than any of our present measures of infant intelligence. If such should prove to be the case, the implications for child placement agencies are indeed far-reaching.

With what is already known about the depressing effect of institutional environment on infant development, especially on their linguistic development, which seems to be most important for their intellectual growth, one wonders whether low test scores on present tests which are heavily weighted with sensori-motor items should be used to retain a child in an institution as possibly "not suitable for adoption," and thus deprive him for a still longer period of the very kinds of stimulation which would be conducive to his verbal and intellectual development.

## REFERENCES

1. ANDERSON, J. E. The limitations of infant and preschool tests in the measurement of intelligence. *J. of Psychol.*, 1939, **8**, 351-379.
2. ANDERSON, L. D. The predictive value of infancy tests in relation to intelligence at five years. *Child Devel.*, 1939, **10**, 203-212.
3. BAYLEY, N. Mental growth during the first three years. *Genet. Psychol. Monog.*, 1933, **14**, 1-92.
4. ———. Consistency and variability in the growth of intelligence from birth to eighteen years. *J. Genet. Psychol.*, 1949, **75**, 165-196.
5. BRADWAY, K. P. Predictive value of Stanford-Binet preschool items. *J. Educ. Psychol.*, 1945, **36**, 1-16.

6. BRODBECK, A., & IRWIN, O. C. The speech behavior of infants without families. *Child Devel.,* 1946, **17,** 145-156.
7. BÜHLER, C., & HETZER, H. Testing Children's Development from Birth to School Age. New York: Farrar & Rinehart, 1935. Pp. 188.
8. CATTELL, P. The Measurement of Intelligence of Infants and Young Children. New York: Psychological Corporation, 1940. Pp. 274.
9. ESCALONA, S. The use of infant tests for predictive purposes. *Bull. Menninger Clinic,* 1950, **14,** 117-128.
10. FISICHELLI, R. M. A Study of the Pre-linguistic Speech Development of Institutionalized Infants. Unpublished doctoral dissertation, Fordham University, 1950. Pp. 116.
11. FURFEY, P. H., & MUEHLENBEIN, J. The validity of infant intelligence tests. *J. Genet. Psychol.,* 1932, **40,** 219-223.
12. GESELL, A. Mental Growth of the Preschool Child. New York: Macmillan, 1925. Pp. 447.
13. GOODENOUGH, F. L. Mental Testing. New York: Rinehart, 1949. Pp. 609.
14. GOODENOUGH, F. L., & MAURER, K. M. The Mental Growth of Children from Two to Fourteen Years. Minneapolis: Univ. Minnesota Press, 1942. Pp. 130.
15. HARRIS, D. B. Intellective functions: Children. *Progress Clin. Psychol.,* 1952, **1,** 26-45.
16. HONZIK, M. P. The constancy of mental test performance during the preschool period. *J. Genet. Psychol.,* 1938, **52,** 285-302.
17. IRWIN, O. C. Research on speech sounds for the first six months of life. *Psych. Bull.,* 1941, **38,** 277-285.
18. ———. Can infants have *IQ*'s? *Psych. Rev.,* 1942, **49,** 69-79.
19. ———. The developmental status of speech sounds of ten feebleminded children. *Child Devel.,* 1942, **13,** 29-39.
20. IRWIN, O. C., & CHEN, H. Speech sound elements during the first year of life; a review of the literature. *J. Speech Disord.,* 1943, **8,** 109-121.
21. ———. Infant speech sounds and intelligence. *J. Speech Disord.,* 1945, **10,** 293-295.
22. KUHLMANN, F. Tests of Mental Development. Minneapolis: Educ. Test Bureau, 1939. Pp. 314.
23. LINFERT, H. E., & HIERHOLZER, H. M. A scale for measuring the mental development of infants during the first year of life. *Stud. Psychol. & Psychiat., Catholic Univ. of America,* 1928, **1,** No. 4. Pp. 34.
24. MCCARTHY, D. Language development in children. In Carmichael, L. (*Ed.*) Manual of Child Psychology. New York: Wiley, 1954. (Pp. 492-630.)
25. MAURER, K. M. Intellectual Status at Maturity as a Criterion for Selecting Items in Preschool Tests. Minneapolis: Univ. Minnesota Press, 1946. Pp. 166.
26. NELSON, V. L., & RICHARDS, T. W. Studies in mental development: I. Performance on Gesell items at six months and its predictive value for performance on mental tests at two and three years. *J. Genet. Psychol.,* 1938, **52,** 303-325.
27. ———. Studies in mental development: III. Performance of twelve-months-old children on the Gesell schedule and its predictive value for mental status at two and three years. *J. Genet. Psychol.,* 1939, **54,** 181-191.
28. SHIRLEY, M. M. The First Two Years: A Study of Twenty-Five Babies. Vol. II: Intellectual Development. Minneapolis: Univ. Minnesota Press, 1933. Pp. 513.

*Department of Psychology*
*Fordham University Graduate School*
*New York City 58*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 211-221.

# CERTAIN EFFECTS OF MESCALINE AND LYSERGIC ACID ON PSYCHOLOGICAL FUNCTIONS*

*The Psychiatric Institute, Columbia University*

CARNEY LANDIS AND JOHS CLAUSEN

## A. INTRODUCTION

During the past few years the psychiatric use of mescaline and lysergic acid (d-LSD-25) with mental patients as an aid for diagnosis and prognosis and as a potential therapeutic measure has received growing attention. The underlying idea is that these drugs have differential effects among normal subjects and the various categories of psychotic patients. Among psychiatric patients, in many instances, these drugs temporarily and markedly aggravate the mental symptomatology, while in normal persons these agents have the effect of disorganizing mental integration.

The vivid hallucinations following the ingestion or injection of mescaline are well-known phenomena. Less well known is the effect of lysergic acid. The administration of this drug may result in hallucinations of any sense modality. Hoch, Cattell, and Pennes (4, 5, 6) related the hallucinations after ingestion of mescaline or lysergic acid to psychopathological characteristics or basic personality trends of the subject. Klüver (8, 9) stressed the sensory changes brought about by mescaline and emphasized the fact that hallucinations occurred in sensory modalities other than vision. Klüver also pointed out that the hallucinatory mechanisms represent general sensory and perceptual characteristics, so indicating that mescaline is an agent by means of which certain fundamental problems in both biology and psychology might be studied.

Such suggestions represent a challenge to psychologists to scrutinize closely the action of these drugs, to determine which psychological functions are modified by the drugs, and which are left intact. We have found no report of the measurement of psychological functions under experimental conditions following the administration of mescaline and lysergic acid in the literature. Any attempt to make a systematic study of these effects entails difficulties since it is to be expected that the vivid hallucinations brought about by the

---

*Received in the Editorial Office on April 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

drugs may have a disturbing effect on most varieties of test performance which might be used.

During the spring of 1951, data on sensory and motor performance were collected by Mr. Stanley Fisher, Mr. David Lascowitz and Miss Lillian Siracuse from seven patients at the Psychiatric Institute who were given mescaline and/or lysergic acid by Drs. P. H. Hoch, J. P. Cattell, and H. H. Pennes. We are indebted to both the psychologists and psychiatrists for these data. Opportunity to report this material did not afford itself until the present time.

## B. METHOD

In the selection of a test battery the criterion employed was that the test performance should be sensitive enough to reflect changes which could clearly be attributed to the chemical agents. The functions measured should be of a rather stable nature, and ones not too likely to be disturbed by hallucinations. Test selection was guided by the experience gained during the study of the effect of psychosurgery on psychophysiological functions (1) and from studies made on the effects of chemical agents (2, 10, 11). The Speed and Endurance of Tapping and the Purdue Pegboard were chosen as measures of psychomotor functions, Critical Flicker Fusion as a measure of retinal and/or cortical sensory function, and Pain Threshold as a test of somaesthetic function.

### 1. *Purdue Pegboard*

The Purdue Pegboard is a finger dexterity test widely used in industry. It provides a measure of the speed of small finger movements with the hands used separately or simultaneously. The subject is required to place small pegs in holes as rapidly as possible with first the right hand, then the left hand, and then using the two simultaneously during successive 30-second periods. The score (Purdue RLB) is the total number of pegs the subject can place during the three periods. He is then required to make assemblies of small parts (pegs, washers, collars) in a designated fashion using both hands as rapidly as possible during a one-minute interval. The score (Furdue Assembly) is the number of pieces assembled during the given interval (Tiffin and Asher, 12).

### 2. *Tapping Speed; Tapping Endurance*

The procedure and the equipment were the same as that utilized in a previous study by Landis and Zubin (10).

### 3. *Critical Flicker Fusion Threshold*

The procedure and the Strobotac equipment were the same as that used by Landis and Zubin (10), (CFF-S), except that three levels of luminance were employed. These were produced by the introduction of Wratten neutral density filters before the test patch which resulted in 7.8, 1.96, and 0.78 millilamberts, test patch luminance.

### 4. *Pain Threshold*

Pain threshold determinations after the method of Hardy, Wolff, and Goodell (3) were included in the test battery, but the thresholds obtained showed such a marked variation that the data cannot be systematically reported.

## C. SUBJECTS

Eleven psychiatric patients (5 male, 6 female; aged 20-54 years; all diagnosed as schizophrenia; and all in the early stages of the disease) constituted the group studied.

Four of the patients served as a control group. and the remaining seven constituted the experimental group. Five patients in the experimental group were given synthetic mescaline sulphate intravenously. One patient received 0.35 gram of the drug while the others received 0.50 gram. Complete data were obtained on all five patients for the Tapping and Purdue Pegboard tests during the period of mescaline intoxication. One patient (No. 5) was unable to do the CFF tests after receiving mescaline. Patients Nos. 2, 3, and 5 experienced hallucinations while Nos. 1 and 4 did not.

Four of the patients included in the mescaline group, and an additional two patients, on whom it had not been possible to obtain data during mescaline, were tested following the ingestion of lysergic acid. One patient (No. 4) was given 0.10 mg of lysergic acid; one (No. 6) was given 0.06 mg first, followed by 0.12 mg 1½ hours later; while the remaining four were given 0.12 mg. In all instances the lysergic acid was given orally. Three (Nos. 2, 3, and 6) of the six patients experienced hallucinations following ingestion of lysergic acid. All three had visual hallucinations; in addition two of the three had auditory and one of the three had olfactory hallucinations.

The patients were given all tests on two or three successive days before mescaline or lysergic acid was administered. Since the number of these test days varied, we have used the last test prior to drug administration as our "pretest" base. The patients were tested about 1½ hours after the drug had been administered, which will be referred to as the "Drug" test. They were again tested about 6 hours after the drug administration which will be referred to as "Post 1" test. Finally they were tested during the morning of the day following the drug treatment, which will be referred to as "Post 2" test. The lysergic acid was administered several days after the mescaline so that the "Post 2" test of the mescaline experiment served as "Pretest" for the lysergic acid experiment for those patients who received both drugs. The patients were tested about 1½ hours after ingestion of lysergic acid ("Drug"), 6 hours after ingestion ("Post 1"), and during the morning of the following day ("Post 2").

The patients in the control group were subjected to the same test battery and to the same test schedule as far as Pretests were concerned. The "Drug," "Post 1," and "Post 2" sessions were one day apart. (The terms "Drug," "Post 1," and "Post 2" sessions, as applied to the control group, really refer to the first, second, and third sessions following the Pretest baseline, since no drug was employed.)

## D. Results

The group averages, standard deviations, and per cent of change based on the pretest level for the various test performances, are given in Table 1. It will be noted that the Pretest score level (except for the Purdue Assembly) of the control group was slightly superior to that of the mescaline or lysergic acid group. Since all three groups were small, the usual tests for statistical significance have not been applied. Our evaluaton of any change in test performance must rest on the size and consistency of the changes as well as on our knowledge of possible experimental errors. Figure 1 gives in a comparative fashion a graphic presentation of the per cent change in score from the pretest score taken as 100 per cent for each test at each test period for the mescaline, lysergic acid, and the control group.

From both Table 1 and Figure 1, it is evident that the mescaline dropped all mean scores for the Purdue Pegboard and Tapping well below those made by the control patients at the equivalent test periods. The loss in efficiency is most marked 1½ hours after injection ("Drug"); it is still evident 6 hours after injection (Post 1), while 24 hours after injection (Post 2) the efficiency is still below that of the control group.

The effect of mescaline on CFF (if any) is unclear with the possible drop of 10 per cent 1½ hours after, and gain of 10 per cent 6 hours after mescaline injection, at the lowest level of luminance. If these changes were significant they probably indicate that mescaline affected only the low intensity CFF thresholds.

TABLE 1
MEANS (M), STANDARD DEVIATIONS ($\sigma$) AND PER CENT OF CHANGE (%) AT EACH TEST PERIOD FOR THE MESCALINE, LYSERGIC ACID, AND CONTROL GROUPS

| | | | Pretest | | | Drug | | | Post 1 | | | Post 2 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Test | Group | N | M | $\sigma$ | % | M | $\sigma$ | % | M | $\sigma$ | % | M | $\sigma$ | % |
| Purdue RLB | Mescaline | 5 | 41.6 | 8.6 | 100 | 37.2 | 4.9 | 89 | 38.0 | 8.2 | 91 | 42.2 | 9.1 | 101 |
| | Lysergic | 6 | 41.7 | 8.0 | 100 | 38.3 | 6.9 | 92 | 38.8 | 5.2 | 93 | 42.5 | 6.4 | 102 |
| | Control | 4 | 43.5 | 5.1 | 100 | 44.0 | 4.9 | 101 | 44.3 | 6.6 | 102 | 44.8 | 5.4 | 103 |
| Purdue Assembly | Mescaline | 5 | 34.4 | 5.2 | 100 | 26.0 | 8.1 | 76 | 28.0 | 9.1 | 81 | 32.4 | 8.5 | 94 |
| | Lyseric | 6 | 30.7 | 8.0 | 100 | 31.7 | 9.1 | 103 | 29.3 | 8.3 | 95 | 34.3 | 8.4 | 112 |
| | Control | 4 | 31.5 | 8.6 | 100 | 35.5 | 5.1 | 113 | 38.0 | 7.0 | 121 | 40.0 | 5.0 | 127 |
| Tapping Speed | Mescaline | 5 | 106.0 | 36.4 | 100 | 95.4 | 35.7 | 90 | 102.4 | 37.0 | 97 | 110.6 | 36.1 | 104 |
| | Lysergic | 6 | 110.8 | 31.5 | 100 | 123.2 | 29.9 | 111 | 112.7 | 19.3 | 102 | 115.5 | 25.3 | 104 |
| | Control | 4 | 125.8 | 34.7 | 100 | 125.0 | 28.4 | 99 | 125.0 | 22.2 | 99 | 134.0 | 29.5 | 106 |
| Tapping Endurance | Mescaline | 5 | 221.8 | 68.8 | 100 | 211.2 | 74.5 | 95 | 205.6 | 70.9 | 93 | 221.2 | 73.1 | 98 |
| | Lysergic | 6 | 228.5 | 49.3 | 100 | 234.5 | 43.6 | 103 | 214.5 | 30.4 | 94 | 245.0 | 53.9 | 107 |
| | Control | 4 | 249.3 | 47.7 | 100 | 256.8 | 43.3 | 103 | 259.5 | 32.7 | 104 | 267.3 | 35.6 | 107 |
| CFF 0.78 ml | Mescaline | 4 | 34.5 | 4.8 | 100 | 31.0 | 7.1 | 90 | 37.8 | 6.4 | 110 | 35.2 | 3.4 | 102 |
| | Lysergic | 6 | 35.5 | 4.3 | 100 | 34.8 | 4.7 | 98 | 35.7 | 5.4 | 101 | 36.1 | 3.6 | 102 |
| | Control | 4 | 37.0 | 2.9 | 100 | 35.6 | 5.7 | 96 | 37.5 | 4.6 | 101 | 35.4 | 1.5 | 96 |
| CFF 1.96 ml | Mescaline | 4 | 36.6 | 3.2 | 100 | 37.7 | 6.4 | 103 | 39.7 | 5.5 | 108 | 37.9 | 4.7 | 103 |
| | Lysergic | 6 | 38.1 | 4.4 | 100 | 36.6 | 4.2 | 96 | 38.0 | 4.1 | 100 | 37.5 | 3.7 | 98 |
| | Control | 4 | 39.0 | 4.5 | 100 | 35.9 | 5.6 | 92 | 36.5 | 1.9 | 94 | 37.0 | 4.2 | 95 |
| CFF 7.8 ml | Mescaline | 4 | 41.8 | 2.8 | 100 | 43.7 | 3.9 | 104 | 42.6 | 3.5 | 102 | 42.9 | 4.5 | 103 |
| | Lysergic | 6 | 40.8 | 4.6 | 100 | 40.9 | 4.3 | 100 | 42.2 | 3.0 | 103 | 42.6 | 3.5 | 104 |
| | Control | 4 | 42.8 | 4.7 | 100 | 45.4 | 4.1 | 106 | 43.0 | 3.0 | 101 | 42.8 | 3.5 | 100 |

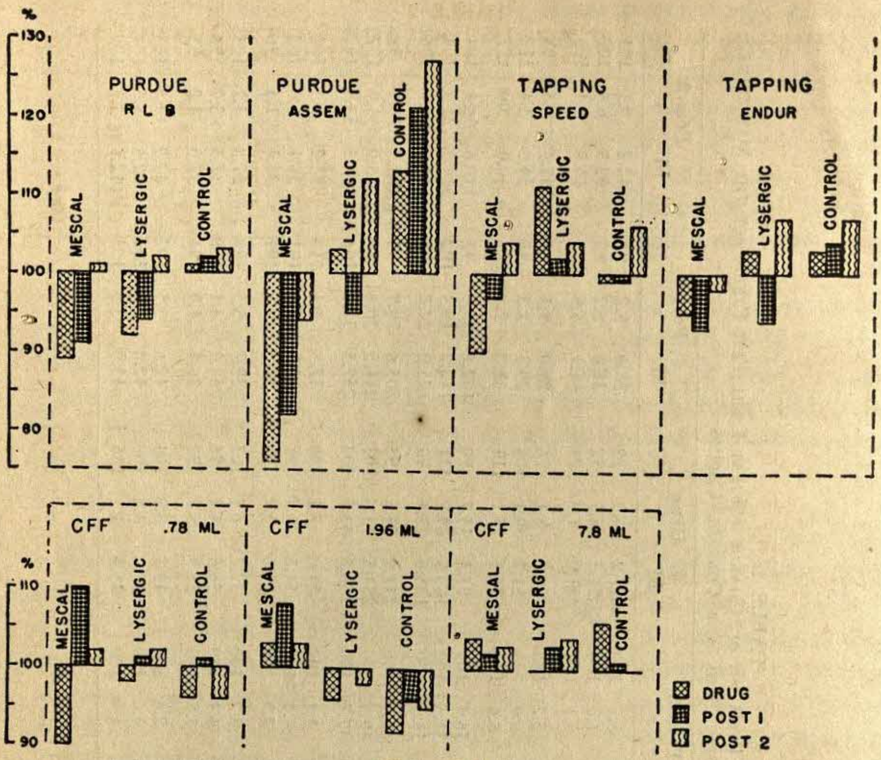


FIGURE 1
TEST SCORES AT EACH TEST PERIOD AS PER CENT OF THE PRE-TEST LEVEL FOR THE MESCALINE, LYSERGIC ACID, AND CONTROL GROUPS

From Table 1 and Figure 1, it will be seen that lysergic acid increased the mean score for Tapping Speed by 11 per cent 1½ hours after ingestion of lysergic acid (Drug) with a return to pretest level 6 hours after ingestion (Post 1). Smaller mean changes of doubtful significance appear in the Tapping Endurance and Purdue Assembly scores. The Purdue RLB score shows a loss of 8 per cent at 1½ hours, 7 per cent at 6 hours, with a return to pretest level at 24 hours. The changes in CFF are minor and probably do not exceed the experimental variability.

It was not to be expected that all patients would react in the same way and to the same degree to the mescaline or to lysergic acid. In Table 2 are given the directions of change (increased or decreased score efficiency) at 1½ hours after injection of mescaline (Drug) and at both 1½ and 6 hours after ingestion of lysergic acid (Drug and Post 1) in contrast to the pretest

TABLE 2

CHANGES IN DIRECTION OF SCORE 1½ HOURS AFTER RECEIVING MESCALINE AND 1½ AND 6 HOURS AFTER INGESTION OF LYSERGIC ACID

| Patient Number | Dosage | Hallucinations | Purdue RLB | Purdue Assembly | Tapping Speed | Tapping Endurance | CFF—.78 ml | CFF—1.96 ml | CFF—7.8 ml |
|---|---|---|---|---|---|---|---|---|---|
| *Mescaline*—1½ hours | | | | | | | | | |
| 1 | .50 gr | — | — | — | — | + | + | — | — |
| 2 | .50 gr | + | — | — | — | — | — | + | + |
| 3 | .35 gr | + | — | — | — | + | — | + | + |
| 4 | .50 gr | — | — | — | + | — | — | + | + |
| 5 | .50 gr | + | 0 | — | — | — | x | x | x |
| *Lysergic Acid*—1½ hours | | | | | | | | | |
| 1 | .12 mg | + | — | 0 | + | + | + | — | — |
| 2 | .12 mg | + | — | + | + | — | + | — | + |
| 3 | .12 mg | + | 0 | + | 0 | + | — | — | — |
| 4 | .10 mg | — | — | 0 | + | — | + | — | + |
| 6 | .06 mg | + | — | — | + | + | — | — | + |
| 7 | .12 mg | — | — | + | + | + | — | — | — |
| *Lysergic Acid*—6 hours | | | | | | | | | |
| 1 | .12 mg | ? | + | + | — | — | + | + | + |
| 2 | .12 mg | + | — | — | — | — | — | + | + |
| 3 | .12 mg | + | + | — | + | — | + | — | + |
| 4 | .10 mg | — | 0 | — | — | — | + | + | — |
| 6 | .18 mg | + | — | — | — | — | + | + | + |
| 7 | .12 mg | — | + | 0 | — | — | + | + | + |
| *No Hallucinations* | | | | | | | | | |
| 1 | Mescaline (.50 gr) | — | — | — | — | + | + | — | — |
| 4 | Mescaline (.50 gr) | — | — | — | + | — | — | + | + |
| 4 | Lysergic Acid (.10 mg) | — | 0 | — | — | — | + | + | + |
| 7 | Lysergic Acid (.12 mg) | — | + | 0 | — | — | + | + | + |
| *Hallucinations* | | | | | | | | | |
| 2 | Mescaline (.50 gr) | + | — | — | — | — | — | + | + |
| 2 | Lysergic Acid (.12 mg) | + | — | — | — | — | — | + | + |
| 3 | Mescaline (.35 gr) | + | — | — | — | + | — | + | + |
| 3 | Lysergic Acid (.12 mg) | + | + | — | + | — | + | — | + |
| 5 | Mescaline (.50 gr) | + | 0 | — | — | — | x | x | x |
| 6 | Lysergic Acid (.06 mg) | + | — | — | — | — | + | + | + |

scores. Since lysergic acid was given by mouth and since Table 1 showed that there was still some evidence of changed scores at Post 1, we have made the additional entries for lysergic acid changes at 6 hours after ingestion.

Inspection of Table 2 shows several interesting points. No patient had an increased Purdue RLB score 1½ hours after either drug, while all patients had a lowered score on Purdue Assembly following mescaline. Lysergic acid gave mixed individual results at 1½ hours and 5 out of 6 scores decreased at the end of six hours. Mescaline usually reduced both Tapping Scores at 1½ hours while lysergic acid usually was followed by an increased score at 1½ hours and a decrease at 6 hours. The CFF changes at 1½ hours are, with the exception of CFF 1.96 ml following lysergic acid, of mixed direction. (The uniform drop at 1.96 ml with lysergic acid at 1½ hours is not convincing in that the amount of drop was in each case small and well within the range of variation shown by the control patients.) However, after six hours almost all CFF scores are increased. We mentioned earlier that the measurements made of pain threshold were so variable that they could not be reported. The CFF measures are also variable and not very conclusive. Hence neither measure of sensory thresholds gave really satisfactory data. On the basis of more recent and as yet incomplete studies, we are inclined to attribute this difficulty in obtaining sensory thresholds to faulty methodology employed in the collection of these data rather than to variability per se.

The effect of these drugs on motor performance may be summarized in the following fashion:

The Purdue Pegboard scores indicate that both mescaline and lysergic acid decrease the efficiency and/or speed of this rather complicated, coördinated motor response. The mescaline acts more profoundly and rapidly and may persist in effect for as much as 24 hours.

Mescaline interferes with both speed and endurance of simple tapping, the effects continuing as long as 6 hours. With the exception of Purdue RLB, lysergic acid tends to act as an immediate stimulant (1½-hour test) with a decrement in most motor scores in most individuals after 6 hours. The effect of either agent is still evident on tapping scores after 24 hours.

In Table 2 we have noted whether or not the drug produced hallucinations. It will be seen that mescaline at the dosage given produced hallucinations in patients Nos. 2, 3, and 5 while lysergic acid did so in patients Nos. 2, 3, and 6. From the inspection of Table 2 in which we commented about lysergic acid, it seems to show a depressant effect on motor test scores and an increased CFF score 6 hours after ingestion. The last two sections of Table 2 are regrouped on the basis of whether or not hallucinations occurred using the mescaline 1½-hour entries and lysergic acid 6-hour entries. This regrouping provides a clearer picture. All patients who had hallucinations (except No. 3) show decreased motor test scores ("0" entries are properly decreases since some increase with learning would be expected to take place). On the CFF measures all patients (except No. 3) who had hallucinations showed an increased score on the 1.96 and 7.8 ml levels. Patient No. 3

had hallucinations following both mescaline and lysergic acid but did not follow completely the pattern of performance of the other hallucinated patients.

The pain threshold determinations were extremely variable, both with regard to any one subject and between the various patients. We were not able to detect any systematic trend in the data. Particularly confusing was the fact that some of the obtained thresholds were completely out of the standard range of variability. We are inclined to attribute this excessive variability of threshold following the two drugs to inherent difficulties in the particular technique, rather than to any real unusual variability of threshold.

## E. DISCUSSION

These experiments demonstrate that controlled objective psychological experiments can be performed when an observer is under the influence of mescaline or of lysergic acid. In other words, neither the "disintegration of personality," the various hallucinations, the existing psychopathology, nor any combination of these, make it uniformly impossible for these patients to coöperate and to provide reasonably stable measures on certain specific tests.

The effect of the injected mescaline was shown by the tests employed 1½ hours after the injection and was still apparent 24 hours after injection so far as the tapping test scores were concerned. When the mescaline produced hallucinations the test scores were, in general, changed more markedly. In the case of patient No. 5, the visual hallucinations were so pronounced that he could not provide satisfactory CFF thresholds although he could do the motor tasks.

The ingestion of lysergic acid, in general, started hallucinations within 1½ hours after taking, and during this initial period tended to produce some increase in the motor test scores. By 6 hours after ingestion, almost all motor test scores were reduced especially if the patient had experienced hallucinations. The CFF threshold was usually increased for the highest brightness level.

That different individuals are affected differentially by different drugs, and that different dosage levels of drugs produce different effects is well known. Considering these points, as well as the limited number of patients involved in this study, the general consistency of results is rather remarkable.

## F. SUMMARY

Patients in the early stage of schizophrenia were used as subjects in a tapping test, in the determination of the critical flicker-fusion threshold and in the Purdue Pegboard test. Five of these patients received synthetic mesca-

line sulphate intravenously, six ingested lysergic acid (d-LSD-25), and four served as a control. The tests were repeated several days prior to the administration of the drug, and 1½ hours, 6 hours, and 24 hours following the administration.

1. In spite of the marked effect of mescaline and lysergic acid on sensation, perception, and general symptomatology of the mental patient, it was possible to measure quantitatively the changes in efficiency reflected by the objective test scores.

2. If the drug induced hallucinatory experience, the changes in test scores tended to be more consistent.

3. Injection of mescaline resulted in a marked mean drop for the group in Tapping and Purdue Pegboard scores 1½ hours following injection which decrease was still apparent 24 hours after injection in the Purdue Assembly and Tapping Endurance scores.

4. Ingestion of lysergic acid (d-LSD-25) produced visual hallucinations in three of six patients. The group mean Purdue RLB score was decreased and the other three motor test scores somewhat increased 1½ hours after ingestion. Six hours after ingestion, except for tapping speed, all group means for the psychomotor tests were reduced, and the scores for the two higher levels of brightness at which CFF was determined were somewhat increased.

5. Analysis of the change in direction of test score performance for each individual concerned showed that when the patient was hallucinated by the drug the individual test results were fairly uniform. Mescaline generally resulted in decreased motor test and lower CFF scores 1½ hours after injection. Lysergic acid usually produced decreased motor test scores and increased CFF thresholds for the highest level of luminance six hours after ingestion.

## REFERENCES

1. COLUMBIA-GREYSTONE ASSOCIATES (F. A. METTLER, *Ed.*). Selective Partial Ablation of the Frontal Cortex. New York: Hoeber, 1949.
2. CRANSTON, R. E., ZUBIN, J., & LANDIS, C. The effect of small doses of thonzylamine, dexedrine, and phenobarbital on test performance and self-ratings of subjective states. *J. of Psychol.*, 1952, **33**, 209-215.
3. HARDY, J. D., WOLFF, H. G., & GOODELL, H. Studies on pain: Discrimination of differences in intensity of a pain stimulus as basis of a scale of pain intensity. *J. Clin. Invest.*, 1947, **26**, 1152-1158.
4. HOCH, P. H. Experimentally produced psychoses. *Amer. J. Psychiat.*, 1951, **107**, 607-611.
5. HOCH, P. H., CATTELL, J. P., & PENNES, H. H. Effects of mescaline and lysergic acid (d-LSD-25). *Amer. J. Psychiat.*, 1952, **108**, 579-584.

6. ———. Effects of drugs; theoretical considerations from a psychological viewpoint. *Amer. J. Psychiat.*, 1952, **108**, 585-589.

7. KING, H. E., & CLAUSEN, J. Psychophysiology. Chapter 13 in F. A. Mettler (*Ed.*), Psychosurgical Problems. Philadelphia: Blakiston, 1952.

8. KLÜVER, H. Mescal: The "Divine" Plant and Its Psychological Effects. London: Kegan Paul, Trench, Trubner, 1928.

9. ———. Mechanisms of Hallucinations. Chapter 10 in Q. McNemar and M. Merrill (*Eds.*), *Studies in Personality.* New York: McGraw-Hill, 1942.

10. LANDIS, C., & ZUBIN, J. The effect of thonzylamine hydrochloride and phenobarbital sodium on certain psychological functions. *J. of Psychol.*, 1951, **31**, 181-200.

11. ———. The alleged sedative effect of thonzylamine hydrochloride (neohetramine). *J. Lab. Clin. Med.*, 1951, **38**, 873-880.

12. TIFFIN, J., & ASHER, E. J. The Purdue Pegboard: Norms and studies of reliability and validity. *J. Appl. Psychol.*, 1948, **32**, 234-237.

*The Psychiatric Institute*
*722 West 168th Street*
*New York City 32*

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] "[illegible]" [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] The [illegible] relative [illegible]

[illegible] 1937, 22, 71-[illegible]

[illegible]

[illegible]

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 223-234.

# IMPRESSIONS OF INFLUENTIAL MEMBERS OF DISCUSSION GROUPS*[1]

*Center for International Studies, Massachusetts Institute of Technology*

HOWARD V. PERLMUTTER

## A. INTRODUCTION

The researches of Asch (1, 2), Haire and Grunes (3), and Kelley (6), have indicated the effects of specific traits on the total impression of a personality. Although they were not explicit in this respect, their concern was primarily with the formation of impressions of individuals whose *direct influence over the perceiver was minimal.* It is clear, however, that in interpersonal relationships, some individuals wield more *influence* than others, and that this property must make a difference to the impression of that person. Two aspects of this interpersonal dimension are termed herewith influence over perceiver and perceived influence over others. Where the influence-over-perceiver of an individual is defined as "the degree to which an individual is perceived as capable of causing satisfaction or dissatisfaction in the perceiver by approval or disapproval," while perceived influence-over-others is defined as "the degree to which an individual is perceived as having influenced other individuals." These definitions are arbitrary to be sure, but their formulation does guide the development of scales for their measurement.[2] This is in general agreement with previous researches on power by Lippitt, Polansky, Redl, Rosen, *et al.* (7).

Many topics then emerge as worthy of investigation: what relations exist between the influence of a person and the perceived complexity of his personality? What is the proportion of desirable and undesirable traits assigned to him?

This present study aims at furnishing some preliminary data relevant to the determinants of trait-descriptions of influential members of a face-to-face group. Our first hypothesis was based on the assumption that the assign-

---

*Received in the Editorial Office on April 17, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This is one of a series of studies connecting perception of power and influence variables with perception of national character images, being conducted at the Center for International Studies, International Communications Program.

[2] Elsewhere Heider distinguished as another property of an individual what is termed here "valence": the attractiveness or desirability of a person (4). Previous research has considered the effects of valence on perception of an individual's actions (5).

ment of a trait to an individual represents an effort by the perceiver to *understand-by-categorizing* (make cognitively clear) the behavior of the influential member. We would expect, then, a greater effort at description of high-influence individuals as compared with low-influence individuals. Hypothesis I was: *The greater the perceived influence of an individual, the greater the total number of different traits will be assigned to him* (e.g., the more complex a person he will seem).

With respect to the kinds of traits assigned we pose Hypothesis II: *The greater the perceived influence of an individual, the more desirable will be the traits assigned him.*[3] This hypothesis was based on the assumption that it is more desirable to see members who can cause satisfaction or dissatisfaction as friendly, rather than unfriendly. It is a derivation from Heider's theory of social perception (4).

## B. PROCEDURE

The subjects were 46 adult members of an evening course in Human Relations at Boston University. The members had been assigned randomly to four groups of 10, 11, 12, and 13 respectively, and had worked as a group on problems of human relations for about 11 weeks. Leadership was rotated at each session so that almost all members of the group had an opportunity to lead the group discussion and to serve as observer. The latter rôle limited participation greatly; the former required active participation.[4]

## C. METHOD

At the end of the eleventh session all subjects were given a series of questionnaires and asked to rank all their fellow-members in the following way: (*a*) The member whose ideas and suggestions most readily (and least readily) were accepted by other members of the group. (*b*) The members whose agreement and disagreement is most (and least) important to the rater. (*c*) The members of the group they would most (and least) prefer working with.

*Dimension 1* is designed to determine those members who are perceived actually to have influenced *the other members* most or least.

---

[3] It is important to note that these hypotheses were developed for and are probably limited to a face-to-face-group situation wherein no individual had intrinsically higher status than another.

[4] Although we cannot assume that participation was controlled in this way, this requirement did insure that all members were brought into the focus of the group during at least one two-hour session. Nevertheless, the rôle of "leader" was interpreted differently by each member. Some played a directive rôle; others led nondirectively.

*Dimension 2* is designed to tap the amount of *influence-over-the-perceiver* attributed to the members of the group. This dimension is most relevant to the above hypotheses.

*Dimension 3* is designed to tap the amount of perceived attractiveness (in a task-oriented situation) of the members of the group.

Finally the subjects were asked to "pick from a list of traits those which characterized best each member of the group," using as many trait adjectives as they wished.

## D. RESULTS

### 1. *Relations between Influence and Valence*

Each member had been rated along three dimensions: those individuals with: high and low (*a*) actual influence-over-others, (*b*) attractiveness, (*c*) influence-over-perceiver. Then a ranking was assigned for each member on Dimension 1 by subtracting from the total number of first or second choices he received as "persons whose ideas and suggestions are most readily accepted" by other members of the group, the total number of first or second choices as "the individual whose ideas and suggestions are least readily accepted" by other members of the group. A high ranking indicated, then, a proportionately greater number of votes as a high actual influencer. Rankings for Dimensions 2 and 3 were determined in exactly the same fashion. For Dimension 3, the total first or second choices for "member of the group they would least prefer working with" were subtracted from total first or second choices as "member most preferable to work with."

Finally, for Dimension 3, the total first and second choices received by each member as the individual "whose agreement and disagreement with what you say is *least* important" is subtracted from the total first or second choice for each member as the individual "whose agreement or disagreement is *most* important." Dimension 1 is called actual-influence-over-others, denoted hereafter as $I_{others}$; Dimension 2, valence (val.); and Dimension 3, influence-over-perceiver ($I_{perceiver}$).

Table 1 indicates the intercorrelation between rankings of group members on the three dimensions. In all groups there is a positive relation among perceived actual-influence-over-others, valence, and influence-over-perceiver. In Group *A* the relationships between valence and actual influence, or influence-over-perceiver are not significant. In Group *B* these relationships are significant at the 5 per cent level; in the other groups there is a higher degree of correspondence between the two influence measurements and valence for three of the four groups. There is a strong tendency then to see the well-

TABLE 1

SUMMARY OF DATA: INTERCORRELATIONS OF RANKINGS OF PERCEIVED INFLUENCE-OVER-OTHERS ($I_{others}$), VALENCE (VAL), AND INFLUENCE-OVER-PERCEIVER ($I_{perceiver}$) FOR GROUP A-D

| | *N* | | Valence A | Valence B | Valence C | Valence D | $I_{perceiver}$ A | $I_{perceiver}$ B | $I_{perceiver}$ C | $I_{perceiver}$ D |
|---|---|---|---|---|---|---|---|---|---|---|
| $I_{others}$ | (11) | A | +.35 | | | | +.87* | | | |
| | (12) | B | | +.74* | | | | +.57† | | |
| | (11) | C | | | +.86* | | | | +.90* | |
| | (13) | D | | | | +.85* | | | | +.67* |
| Val | | A | | | | | +.31 | | | |
| | | B | | | | | | +.58† | | |
| | | C | | | | | | | +.88* | |
| | | D | | | | | | | | +.74* |

*1% level of significance.
†5% level of significance.

liked individuals as influential in group process and having greater influence over the perceiver. The correlation between actual influence and influence-over-perceiver rankings of members of all four groups are all significant at least at the 5 per cent level. On the other hand, the results suggest that there is some variability, however small, in the degree to which individuals who are perceived as having high influence over others are also perceived as having high influence over the perceiver.

### 2. *Influence and Total Traits Assigned*

Table 2 summarizes the correlations among influence, valence, and total traits assigned of actual influence, valence, and influence-over-perceiver with the total number of traits assigned to each member. The correlations are consistently highest with influence-over-perceiver ratings, and are lowest with valence.

TABLE 2

SUMMARY OF DATA: RELATIONSHIP OF $I_{others}$, VALENCE, AND $I_{perceiver}$ WITH TOTAL NUMBER OF TRAITS ASSIGNED EACH GROUP MEMBER

| *N* | Group | $I_{others}$—No. Traits | Valence | $I_{perceiver}$ |
|---|---|---|---|---|
| 10 | A | +.64† | +.03 | +.82* |
| 13 | B | —.16 | —.26 | +.32 |
| 11 | C | +.13 | +.18 | +.47 |
| 13 | D | +.43 | +.52‡ | +.77* |

*1% level of significance for one-tailed test.
†2% level of significance.
‡5% level of significance.

But in Group *B*, one particular member ranked last in actual influence, valence, and influence-over-perceiver, but ranked first in total traits assigned. Of the 53 traits assigned to him, only 18 were desirable, 22 were undesirable, and 23 that were neutral could be construed as undesirable in their contents. Rankwise, this individual was second in the number of negative traits assigned. To some extent, this accounts for the relatively low correlations in Group *B*, as compared to other groups where no such controversial person was present. It suggests too, that the very obnoxious member with little influence may be assigned many traits, the majority of which are unpleasant.

The data then tend to support Hypothesis I for influence-over-perceiver more strongly than for influence-over-others. No strong relationship exists between valence and number of traits assigned.

### 3. *Influence, Valence, and Desirability of Traits*

It was possible as well to divide the trait list into three groups:

*a. Desirable traits:* cheerful, adventurous, self-sufficient, serious, good-natured, broad interests, mature, easy-going, adaptable, coöperative, trustful, imaginative, frank, attentive, democratic, friendly, poised;

*b. Undesirable traits:* depressed, worrying, dependent, frivolous, narrow interests, easily upset, aggressive, passive, silent, rigid, unshakeable, obstructive, unimaginative, autocratic, aloof;

*c. Neutral traits:* (as seen by 10 independent judges from the experimental groups, and observers of group process) placid, talkative, unshakeable poise, motherly, reserved, dominant, cautious.

Thus the relationship between influence-over-others, influence-over-perceiver, and desirability of traits assigned was determined. If Hypothesis II is to be supported, we should expect high positive correlations between influence and desirability of traits.

Table 3 shows the correlations between rankings on $I_{others}$, $I_{perceiver}$, and valence with the number of desirable and undesirable traits assigned to that individual. In addition, a *Rank Difference Score* was calculated for each member in the following way: the number of undesirable traits were subtracted from the number of desirable traits. The difference scores were given relative rankings. In this way the assignment of undesirable traits reduced the ranked desirability of the individual.

The correlations were relatively high with Rank Difference Scores for all four groups. Furthermore, influence-over-perceiver rankings correlated positively at the .01 level of significance with the number of desirable (D) traits assigned, and negatively in Group *C* with the undesirable (U) traits

## TABLE 3

SUMMARY OF CORRELATIONS: RANK $I_{others}$, VAL, $I_{perceiver}$, AND RANK NUMBER OF DESIRABLE (D) AND UNDESIRABLE (U) TRAITS AND RANK DIFFERENCE SCORE (R.D.)

| | *D* | | | | *U* | | | | *R.D.* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | A | B | C | D | A | B | C | D |
| $I_{others}$ | +.82* | +.56† | +.63† | +.76* | —.37 | —.65† | —.77* | —.74* | +.70† | +.85* | +.85* | +.83* |
| Val | +.08 | +.46 | +.59† | +.86* | —.35 | —.68* | —.07 | —.72* | +.11 | +.80* | +.76* | +.94* |
| $I_{perceiver}$ | +.80* | +.88* | +.77* | +.84* | —.49 | —.45 | —.66† | —.41 | +.72† | +.72* | +.91* | +.76* |

*1% level of significance.
†5% level of significance.

assigned. The negative correlations between $I_{others}$ and desirable and undesirable traits are almost as high. The Rank Difference Score seems most closely connected with both actual influence-over-others and influence-over-perceiver.

Table 3 indicates also the relation between valence and the number of desirable and undesirable traits. Here one would expect high positive correlations with desirable traits and negative correlations with undesirable traits. In two of the groups, significant positive correlations between valence ranking and number of desired traits were noted, and in three of the four groups the Rank Difference Score correlated significantly.

### 4. *Trait Assignments of High and Low Influence Members*

Let us now examine in some detail the characteristics of those particular members who received most and least choices as having greatest and least influence over the perceiver. Four individuals were selected from each of the four groups. Two were called high $I_{perceiver}$ members. These were the members picked most frequently as first or second choice for "the member whose agreement or disagreement with your contributions is *most* important." Two were called low $I_{perceiver}$ members. These were members whose agreements and disagreements are "least important." Table 4 shows the average total choices (first and second) for the high and low $I_{perceiver}$ individuals for each of the four groups.

TABLE 4
AVERAGE NUMBER OF CHOICES FOR HIGH AND LOW INFLUENCE ($I_{perceiver}$) MEMBERS FOR GROUPS A-D

| | A | | B | | C | | D | |
|---|---|---|---|---|---|---|---|---|
| | M.I.* | L.I.** | M.I. | L.I. | M.I. | L.I. | M.I. | L.I. |
| High Influence | 6.5 | 1 | 4.5 | 0 | 4 | 0.5 | 5 | 0 |
| Low Influence | 0 | 4.5 | 0 | 3.5 | 0.5 | 3.5 | 0.5 | 4.5 |

*M.I.: agreement or disagreement *most* important.
**L.I.: agreement or disagreement *least* important.

The total number of traits (irrespective of content) assigned to eight high influence and eight low influence individuals is shown in Table 5. The high $I_{perceiver}$ individuals clearly receive significantly more desirable traits, a few more neutral, and many fewer negative traits, as we have indicated before.

The low $I_{perceiver}$ members receive more undesirable traits than desirable ones, but the difference is not so great. On the average, low $I_{perceiver}$ group members appear more ambivalent about low influence members.

TABLE 5
AVERAGE NUMBER OF TRAITS ASSIGNED PER PERSON FOR HIGH AND LOW INFLUENCE ($I_{perceiver}$) INDIVIDUALS

| | Positive traits | Negative traits | Neutral | Total |
|---|---|---|---|---|
| High influence | 32.0 | 3.8 | 7.2 | 43.0 |
| Low influence | 12.5 | 16.1 | 4.6 | 33.2 |

### 5. *Traits Most Frequently Assigned High and Low Influence Individuals*

From Table 6 it is clear that low $I_{perceiver}$ individuals are seen as "silent," "passive," and "autocratic" more often than high $I_{perceiver}$ individuals. On the other hand, these high influence individuals are seen as "mature," "coöperative," and "democratic" more frequently. This is no doubt in conformity with the avowed norms of discussion groups "aimed at functioning democratically."

TABLE 6
RANKING OF TRAITS FOR EIGHT HIGH AND EIGHT LOW INFLUENCE ($I_{perceiver}$) MEMBERS OF FOUR DISCUSSION GROUPS ($N = 35$)

| High influence | Ranking | Corresponding low influence ranking | Low influence | Ranking | Corresponding high influence ranking |
|---|---|---|---|---|---|
| Nature | 1 | 15 | Silent | 1 | 24 |
| Coöperative | 2 | 9 | Passive | 2 | 35 |
| Frank | 3 | 18 | Serious | 3 | 9 |
| Cheerful | 3 | 5 | Autocratic | 3 | 26 |
| Democratic | 5 | 22 | Good-natured | 3 | 6 |
| Self-Sufficient | 6 | 14 | Cheerful | 6 | 3 |
| Good-natured | 6 | 3 | Attentive | 7 | 7 |
| Broad interests | 8 | 18 | Reserved | 8 | 26 |
| Attentive | 9 | 7 | Cautious | 9 | 14 |
| Talkative | 9 | 13 | Submissive | 9 | 35 |
| Serious | 9 | 3 | Coöperative | 9 | 2 |
| Adaptable | 12 | 22 | Placid | 9 | 17 |
| Dominant | 13 | 22 | Talkative | 13 | 9 |
| Imaginative | 14 | 27 | Self-sufficient | 14 | 6 |
| Cautious | 14 | 9 | Easily upset | 14 | 24 |
| | | | Narrow interests | 14 | 26 |

But Table 6 shows little difference in the relative ranking of the trait "talkative" which suggests that while some low influence individuals are seen as "silent or introspective," a considerable number are seen as "talkative." We suspect that the significance of the "talkative" trait-characterization depends (as is the case with all other traits) on the constellation of other traits used. In these groups, individuals were called "talkative and democratic" if they tried to stimulate group discussion, encourage a group decision, and suggest new ideas with appropriate modesty. On the other hand, the "talkative-autocratic" individual was probably connected with such actions as: directive attempts at manipulating another person's opinion, monopolizing the discussion with polemic discourses, etc.

Nevertheless, some of the traits assigned both to high and low influence individuals are clearly desirable and neutral. The relative rankings of traits such as "cheerful," and "attentive" are similar for both groups.

In general, though, adding the above characteristics of the low influence individual to the fact that low influence individuals receive more assignments of traits such as "narrow interests," "easily upset," "cautious," and "reserved," one receives the picture of the low influence individual as relatively unstable and rigid.

## D. Discussion of Results

### 1. *Influence and Trait Assignment*

Our two main results were interrelated: the influential members of a discussion group are assigned more traits per person than the less influential members of the group when influence is measured in terms of the direct influence-over-the-perceiver rather than influence-over-others (despite a relatively high intercorrelation between rankings for influence-over-others and influence-over-perceiver). Why should influence-over-perceiver be related to total trait assignment more closely than actual influence-over-others rating? Direct power or influence by a person over the perceiver induces a greater effort to control by describing and/or a greater awareness of the source of influence. To put it another way, some high influence-over-others individuals are not necessarily influential to the perceiver and hence not linked as directly to the perceiver (and to Hypothesis I and the assumptions underlying its formulation).

The Rank Difference Scores for each individual (an index of excess of desirable traits over undesirable traits) correlate highly with actual influence over others, valence, and influence over the perceiver. It appears, then that Hypothesis II received adequate support.

The objection can arise, however, that if high influence members receive more total traits they should receive more desirable traits and, logically, *more* undesirable and neutral traits. But the results indicated that while high influence individuals receive more desirable traits they also receive *fewer* undesirable traits, and approximately the same number of neutral traits as low influence persons. Furthermore, the intercorrelations of valence and influence are sufficiently high to support the contention that it is more likely that high influence individuals in a discussion group will be seen as friendly and attractive.

### 2. *Participation and Influence*

It is possible that the high participators were high influence individuals. While no measurements were made on total participation during group process, there is some good qualitative evidence relating to the question: do high influence individuals furnish more information about themselves by talking more frequently than others?

In our discussion groups, wherein members were oriented toward *resisting direct attempts* at manipulating attitudes and decisions of other members, the "talkers" were not seen as most influential. The "good group members," e.g., those with influence, aimed at clarifying ideas for other members, and suggesting, relatively infrequently, their own ideas. They refrained from directive attempts at influence and dominating the discussion. In the post-meeting questionnaires, in no case were individuals chosen by group members as "the individual who talks most" also chosen as the individual "who contributes most," and by implication "the individual who influences most."

The high influence members were usually associated with the functions of "clarifier," "summarizer," "harmonizer," while the low influence members were seen as connected with the functions of "axe-grinder" and "monopolizer." As was noted before, the members chosen as the individuals who "talk least" in some cases were low influence members.

Generally, for these groups it was more likely that the high influence individual participated frequently in accepted, "non-directive" ways, remaining cognizant of other members' needs to contribute to the group discussion.

### 3. *Further Notes on Influence and Perception*

Let us examine further our preliminary formulation: the greater the perceived influence of an individual the greater the forces that arise within the perceiver to control that influence. We consider this effort to control one of many devices employed to resolve the state of disequilibrium postulated as existing between perceiver and the perceived source of influence over him. In this condition, one individual governs in some unilateral way the behavior of the perceiver by his capacity to reward or punish the perceiver for a series of behaviors not yet performed and not cognitively clear. Whether the high influence individual rewards or punishes more frequently, no doubt is of some consequence. But from the point of view of the perceiver, making predictable the behavior of the high influence individual is requisite to achieving more rewards than punishments from him. It would follow then that a greater effort would be made to describe and categorize the properties of the high influence individuals rather than the low influence individuals. The process of trait assignment represents a crude but certainly not magical way of reducing the power of a high influence individual to withhold reward or to punish.

We would suggest too that the obnoxious member, although not consciously perceived as a high influence group member, probably exerts a high degree of influence on group process, judging from the efforts to characterize the "monopolizers" mostly with undesirable traits. In this case, however, the intent was not only to express aggression toward this source of frustration, but also to reject his many attempts at influence by labeling the variety of ways he tries to exert his negative influence.

Future experimentation should, however, aim at unraveling the complicated relationships between perceived influence of a source, amount of attention to the source, and information about that source of influence.

## E. Summary and Conclusions

Results from four adult discussion groups of 10, 11, 12, and 13, respectively, tend to support the two hypotheses posed:

I. The greater an individual's perceived capacity to influence the perceiver, the greater the total number of traits will be assigned to him. A weaker relationship was found between perceived actual influence over others and total number of traits assigned.

II. The greater the perceived influence of an individual the more desirable will be the traits assigned to him.

The results then suggest that influence over the perceiver by an individual affects the judgment of that person by either increasing awareness of and/or an increasing need to describe the behavior of the source of influence. But further research is needed to establish the relative importance of perceived influence of an individual, awareness of the influential individual, and information available concerning the influential individual.

## References

1. Asch, S. E. Forming impressions of personality. *J. Abn. & Soc. Psychol.*, 1946, **41**, 258-290.
2. ————. Social Psychology. New York: Prentice-Hall, 1952.
3. Haire, M., & Grunes, W. F. Perceptual defenses: Processes protecting an organized perception of another personality. *Hum. Rel.*, 1950, **3**, 403-412.
4. Heider, F. Attitudes and cognitive organization. *J. of Psychol.*, 1946, **21**, 107-112.
5. Horowitz, M. W., Lyons, J., & Perlmutter, H. V. Induction of forces in discussion groups. *Hum. Rel.*, 1951, **4**, No. 1.

6. Kelley, H. H. The warm-cold variable in first impressions of persons. *J. Pers.*, 1950, **18**, 431-439.

7. Lippitt, R., Polansky, N., Redl, F., & Rosen, B. The Dynamics of Power: A Field Study of Social Influence in Groups of Children. From: Swanson, Newcomb, Hartley, *et al.*, *Readings in Social Psychology*. New York: Holt, 1952.

*Center for International Studies*
*Massachusetts Institute of Technology*
*Cambridge 39, Massachusetts*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 235-250.

# AMPLITUDE AND LATENCY CHARACTERISTICS OF THE CONDITIONED HEART RESPONSE*[1]

*Department of Psychology, University of Connecticut*

D. ZEAMAN, G. DEANE, AND N. WEGNER

## A. THE PROBLEM

Our problem is the experimental one of demonstrating orderly changes in both latency and amplitude measures of the human heart rate during the conditioning of that response. The theoretical significance of this problem stems from several sources. (*a*) There are theories concerning the interrelationships among measures of conditioning (3), and crucial tests of these require simultaneous, trial-by-trial measures of amplitude and latency (8). According to Hull (2), and Skinner (6, p. 112), latency and amplitude measures can be derived best, and perhaps only, from the conditioning of autonomically mediated responses. Since the heart rate is a good example of an autonomic response, its use is therefore recommended. (*b*) The currently wide usage of anxiety as a theoretical construct warrants a detailed examination of the properties of any of the possible indices of anxiety such as heart rate. (*c*) The data from heart rate conditioning bear heavily upon various explanations of the mechanism of conditioning. For example, the finding reported by Notterman, Schoenfeld, and Bersh (4) of a heart rate *CR* markedly dissimilar to the *UR* proves embarrassing to at least one theory, the Stimulus Substitution view, which calls for identical *CR* and *UR*. However, a description of only a small portion of the heart rate *CR* has yet been given. The form of the response being important here, a closer look at the amplitude and time characteristics of the complete *CR* and *UR* is clearly indicated.

The present study makes use of a heart conditioning technique developed recently by Notterman, Schoenfeld, and Bersh (4). It represents a partial replication and extension of their work with special attention given to a beat-by-beat analysis of the response and the addition of a more rigorous pseudo-conditioning control test.

---

*Received in the Editorial Office on April 18, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This is based on Technical Report No. 6 under Contract Nonr-631 (00) between the Office of Naval Research and The University of Connecticut.

## B. SUBJECTS, APPARATUS, AND PROCEDURE

### 1. *Subjects*

The subjects were 28 students at the University of Connecticut ranging in age from 20 to 27 years. These were divided into two groups, a Conditioning Group made up of 14 men and 3 women, and a Pseudoconditioning Group consisting of 9 men and 2 women. All subjects were paid volunteers.

### 2. *Apparatus*

In general the apparatus consisted of an electrocardiograph to measure the heart rate, a source of shock for the *US*, a tone for the *CS*, and electronic timers for controlling the sequence of stimuli.

More specifically, the apparatus included a Garceau Junior Electroencephalograph converted for use as an electrocardiograph and stimulus recording kymograph, a Maico Audiometer for tone source, and a Harvard inductorium as shock source. The durations of and interval between *CS* and *US* were controlled by three Hunter Decade Interval timers. The shock was 13 *V.A.C.* applied across the first two fingers of the subject's left hand by means of copper electrodes covered with electrode jelly. The tone, delivered through earphones, was set at 60 db and 512 *c.p.s.* The subject was seated in a separate cubicle, lined with acoustical celotex and having the dimensions 4¼′ x 4¼′ x 8′. All the apparatus with the exception of the shock electrodes, cardiograph leads, and earphones was outside the subject's cubicle. An electric fan in the wall of the cubicle provided ventilation and an effective sound screen for any cues from the experimenter or apparatus.

### 3. *Procedure*

Parts of our procedure and technique were closely patterned after those of Notterman, Schoenfeld, and Bersh (4). Each subject went through two phases within the same 1½-hour experimental session. For the first phase all subjects were given a series of preconditioning trials consisting of the presentation of 20 1-sec. tones spaced at irregular intervals of one to two minutes. The second phase immediately followed this with 20 further trials of either conditioning or pseudoconditioning. A conditioning trial consisted of a 1-sec. tone followed in six seconds by a 6-sec. shock. During pseudoconditioning the subject received on each trial *either* a 1-sec. tone *or* a 6-sec. shock but not both. The pseudoconditioning trial sequence of tone (*t*) and shock (*s*) was irregular: stttsttttsttttsttsttttt. Intertrial intervals for both conditioning and pseudoconditioning varied irregularly between one and two min., as during pre-conditioning.

Cardiograph records were taken on each trial from about 20 sec. before the tone to 40 sec. after. Resulting cardiograms looked similar to a published sample (4).

The subjects were instructed that they would receive a series of tones and a series of shocks, but the order was not mentioned. The instructions were the same as those used by Notterman, Schoenfeld, and Bersh (4).

## C. RESULTS

### 1. *Plotting the Form of the Response*

In the top half of Figure 1 there are plotted the changes in heart rate produced by the tone alone during the first 20 trials (pre-conditioning) and the changes produced by the tone and shock during the second 20 trials (conditioning). These functions are derived in the following way. On a single trial for one subject approximately 35 heart beat intervals are measured in millimeters from the cardiograph tape. There are 10 beat intervals measured just before the onset of the tone, one beat measured during the tone, all the beats measured between the tone and the shock, all the beats measured during the "shock" period regardless of whether there is a shock or not, and finally 10 beats after the "shock" period. These measurements are averaged over all 17 conditioning subjects for a single trial and then converted to beats per minute. An average is then taken over the first 20 trials to yield the pre-conditioning curve, and another average over the second 20 trials to give the form of the response during the conditioning trials.

### 2. *The Tone UR*

It can be seen that during the pre-conditioning period the average heart rate was about 77 b/m for the 10 beats prior to the tone onset. During the tone the rate dropped about 1½ b/m. Following the tone offset, during the 14 beats after the tone, the rate underwent a compensatory rise almost to the pre-tone level before showing a secondary depression and rise. There was also a suggestion of a tertiary depression and rise in the period marked "post-shock" (no shock in pre-conditioning phase), completing the picture of the human heart's unconditioned response to tone. The response can be described in general as a depression in rate followed by progressively damped oscillation.

### 3. *The CR to Tone*

Attention to the "conditioning" curve in the top half of Figure 1 reveals the change in heart response to tone brought about by its association with shock. The main effect of conditioning is the accentuation of the secondary

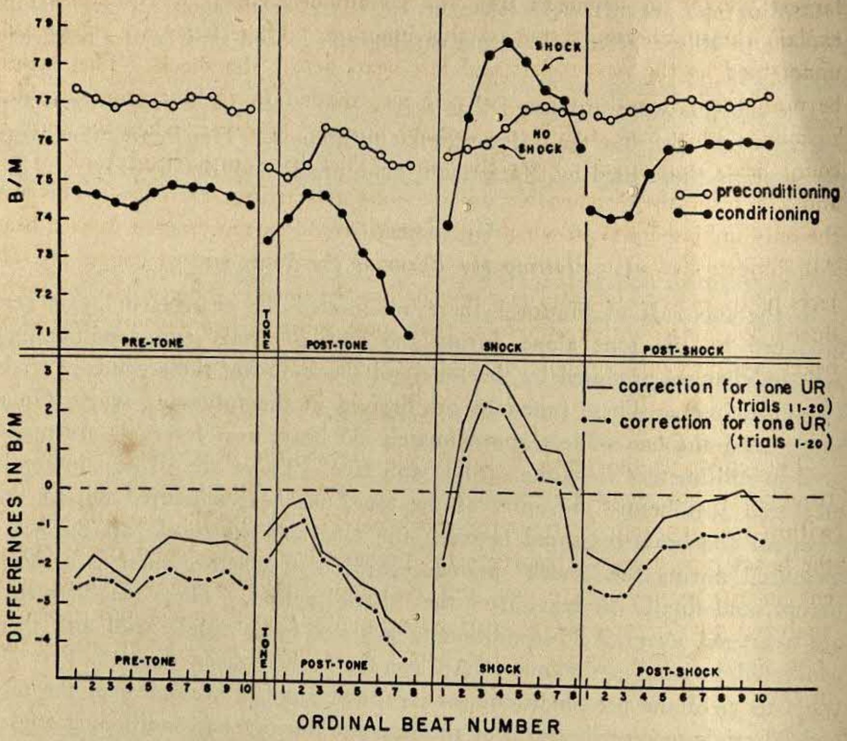


FIGURE 1

THE FORMS OF THE UNCONDITIONED AND CONDITIONED RESPONSES OF THE HEART

The ordinate has the dimension of average heart rate in beats per minute. The abscissa represents the ordinal number of heart beats, counting forward from the tone and shock as successive reference points in the trial cycle.

On the top half of the figure, the unfilled circles show the unconditioned response to the tone during the first 20 trials (pre-conditioning, tone-alone). The filled circles show the conditioned response to the tone during the 8-beat period marked "post-tone," and the unconditioned response to shock during the periods labelled "shock" and "post-shock." Both of the above functions are averages of 17 subjects taken over 20 trials.

The dot-dash line on the bottom half of the figure is a plot of the differences between the two curves in the top half of the figure. The solid line is a similar function using the last 10 pre-conditioning trials as a correction instead of all 20 trials.

depression. The magnitude of this depression from pre-tone level is about 4 b/m at its maximum point just before shock. This is a highly significant effect if we use the small pre-tone variability as a measure of experimental error. The curious fact may be noted that the last two beats before the shock are plotted as ordinal beats 6.5 and 7.5. Since these beats carry the largest weight in demonstrating the conditioned effect, it is necessary to explain the numbering system at this juncture. These beats are most easily understood as the next-to-last and last beats before the shock. They would be numbered 2 and 1 respectively if we counted in the direction of shock to tone. They have a different ordinal number, however, when we start to count beats immediately after the tone. They have an average rank of 6.5 and 7.5 respectively when we use the tone as starting reference. This was the only unbiased way in which all subjects would be represented in each beat. All subjects had at least six beats following the tone, and of course all subjects had a next-to-last and last beat before the shock. The only error introduced by this counting system is that some subjects had one of their sixth beats averaged into not only the sixth beat but also the 6.5th beat.

### 4. *Shock UR*

The effect of the shock is to sharply elevate the heart rate. From the trough brought about by conditioning, the rate is brought up about 7½ b/m within the first four beats of the shock period. While the shock is still on, the heart rate starts its return to the pre-tone level. A second rise is apparent in the post-shock period, and for those subjects whose records were analyzed for 20 post-shock beats a series of damped oscillations appeared. The shock *UR* is similar to the tone *UR* in being a wave-like disturbance in rate.

### 5. *The CR Corrected for Tone UR*

Since our *CS* was by no means a neutral stimulus, an attempt has been made to correct for this difficulty. If the form and magnitude of the tone *UR* have been accurately assessed during the pre-conditioning period, and this response tendency remains constant throughout conditioning, a straightforward correction would consist of a subtraction of the response during the first 20 trials from that measured during the second 20 (conditioning) trials. The bottom half of Figure 1 shows the result of this correction. The dot-dash line shows the difference between the two functions plotted in the top half of Figure 1. Thus corrected the *CR* becomes first a rise in rate over the pre-tone level and then a drop before the shock. It is possible that this initial rise is an artifact. If the tone *UR* has adapted out by the time that conditioning begins, just such a rise would result from this correction procedure. As a check against this possibility, a correction has been made using only the last 10 pre-conditioning trials. If tone *UR* adaptation were occurring, the initial rise effect would disappear. The solid line depicts the result of this correction. The parallelism of the two curves is evidence against the notion that the initial rise is an artifact due to adaptation.

The negative values of the corrected *CR* mean only that the over-all level of heart rate (the tidal level, so to speak) fell progressively throughout the course of the experiment.

### 6. *Form of the Pseudoconditioned Response*

There are three functions instead of two to be presented for the 11 pseudoconditioning subjects: one showing the form of the tone *UR* during the preconditioning phase, one showing the form of the tone response during the pseudoconditioning phase, and another the *UR* to shock.

The dashed line in the top half of Figure 2 shows the *tone* UR averaged over the first 20 (pre-conditioning) trials. It was computed in exactly the same way as the analogous function in Figure 1, and the results differ very little. The first compensatory rise after the tone is slightly higher for the pseudo group than for the conditioning group, but that is all.

The shock *UR* (dot-dash line) also resembles closely that obtained from the conditioning subjects. The fact that the over-all rate level is higher on the shock trials than on the others can be attributed to momentary fluctuation of general level during the pseudoconditioning trials. The shock curve is made up of an average of only five shock trials in comparison to the other two more stable functions containing averages of 10 and 20 trials. One significant feature of the function representing these five shock-alone trials is the absence of any response where the tone would normally be. This can be interpreted as proof of the absence of any extraneous cues from the apparatus or experimenters serving as a warning of approaching shock.

The curve with filled circles is made up of the averages of the five pairs of tone-alone trials during the pseudoconditioning trials. Each pair followed one of the five shock-alone trials. These 10 trials were selected on the grounds that pseudoconditioning would be expected to be greatest here, if anywhere. The unconditioned response to tone is still in evidence, and its form and magnitude are not changed greatly despite its being interspersed among a series of shocks. We consider our pseudoconditioning procedure as more rigorous than that previously employed (4) as control because it more closely approximates the conditioning procedure (tone and shock) without including the *CS-US* pairing feature that defines conditioning.

The difference functions on the lower half of Figure 2 are analogous to those of Figure 1. The dot-dash line expresses the difference between the response to tone during the first 20 trials and the response to tone during

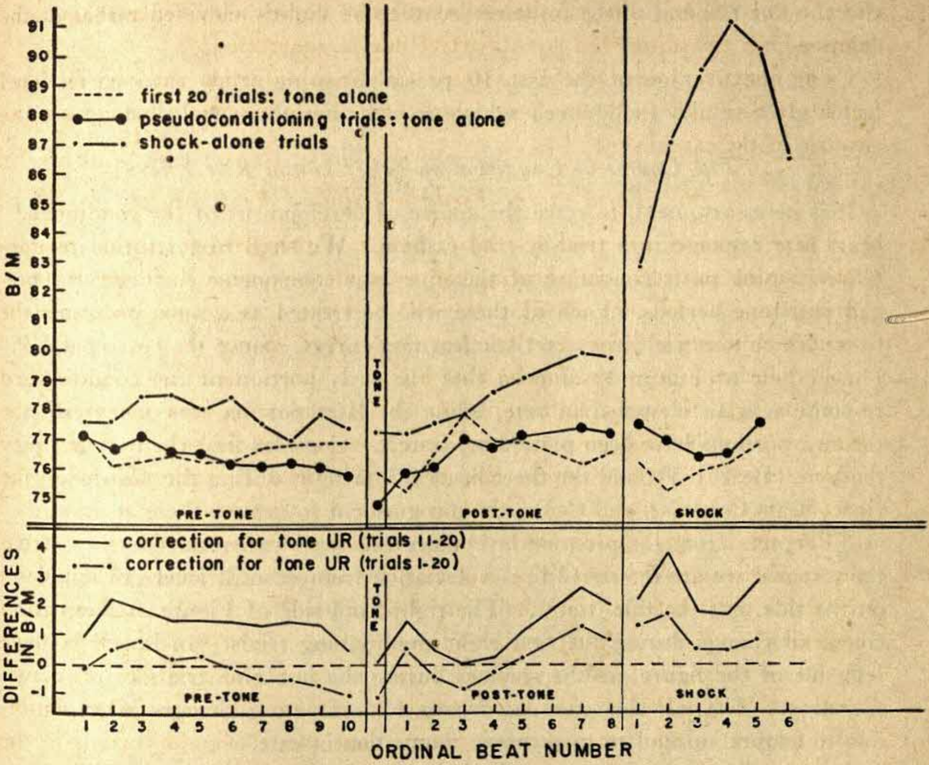


FIGURE 2

THE FORMS OF UNCONDITIONED AND PSEUDOCONDITIONED RESPONSES OF THE HEART: THE ABSCISSA AND ORDINATE ARE THE SAME AS THOSE OF FIGURE 1

For the top half of the figure, the dashed line shows the unconditioned response to the tone during the first 20 (tone-alone) trials. The filled circles show the presumably pseudoconditioned response to the tone during 10 trials of pseudoconditioning. The dot-dash line shows the unconditioned response to shock during the five shock-alone trials of pseudoconditioning.

All the curves in the top half of the figure are based on averages of 11 subjects. The bottom half of the figure shows the difference (dot-dash line) between the first 20 trials (dashed line) and the pseudoconditioning (filled circles) curve of the top half of the figure. The solid line uses the last 10 trials of the first 20 as correction factors.

the second 20. This function fluctuates about the zero difference level for most of the responses. If we compute a mean and standard deviation from the 10 pre-tone values, only one of the points after the tone lies more than two SD units away from the pre-tone mean. A significant pseudoconditioned response is clearly absent. What little tendency there is for a sensitization or pseudoconditioning effect is in a direction opposite to that of the conditioned effect. The tail end of the response seems to be slightly elevated rather than depressed.

Using the average of the last 10 pre-conditioning trials as a correction factor gives results (solid line) which do not alter this conclusion.

### 7. *The Course of Conditioning—the "Initial Rise Effect"*

It is necessary, next, to trace the course of development of the conditioned heart rate response in a trial-by-trial fashion. We shall first fractionate the *CR* into nine parts consisting of the nine beat components during the tone and post-tone periods. Each of these will be treated as a separate response to see which ones yield characteristic learning curves. Since the corrected *CR* (lower half of Figure 1) showed that the early portion of the conditioned response was an elevation in rate, while the later portion was a depression, the two portions have been plotted separately. Consider first the beat during the tone (Beat T-1), and the three beats following it during the post-tone period (Beats C-1, C-2, and C-3). It is convenient to express these beat values as differences from the pre-tone level since this level varies considerably, and the response we are interested in is a deviation from over-all level (the ripples on the tide, not the tide itself). The right-hand side of Figure 3 shows the course of change during the first eight conditioning trials. Included on the left side of the figure are the changes during the last nine trials of pre-conditioning. The left side gives us a basis for estimating changes in response due to adaptation and to momentary fluctuation in rate.

As a detailed example, we can read the C-1 function as follows. The first beat after the tone has a value of about 2.4 b/m on the trial numbered 13. This means that at this stage of pre-conditioning (Trials 12, 13, and 14 averaged) the first post-tone beat was 2.4 b/m slower than the mean of the 10 pre-tone beats. By Trial 19 its value has followed an irregular course downward (adaptation) until conditioning begins. Conditioning produces a fall in the curve beyond that which would be expected by adaptation. This sudden fall (the loss of a slow-down) is responsible for the "initial rise effect" in the form of the corrected *CR* shown in Figure 1. The change brought about by the conditioning procedure is not permanent, however. There is a recovery to the pre-conditioning level by Trial 7, demonstrating that the effect is a temporary one. If the effect were due to progressive adaptation of the tone *UR* it could not show such a recovery. The "initial rise effect" is attributable to the conditioning procedure, but the effect is not stable.

A similar interpretation can be made of the other functions. Beat C-2 in particular follows a course similar to that of C-1.

### 8. *The Course of Conditioning—the "Depression Effect"*

Figure 4 is to be read like the previous figure. Here are plotted the values for the remainder of the post-tone beats, C-4, C-5, C-6, C-6.5, and C-7.5 representing the depressed portion of the CR. The values for these beats are in general less variable during the last part of pre-conditioning and quite regular during the conditioning trials. The functions are readily interpretable as learning curves. The learning reaches a plateau at about the seventh trial, following which there is a gradual decline in the magnitude of the CR out to Trial 20. The response does not disappear, however, as is shown by the mean values for the last five trials. This decline in the strength of the *CR* may be related to the adaptation of the shock *UR*.

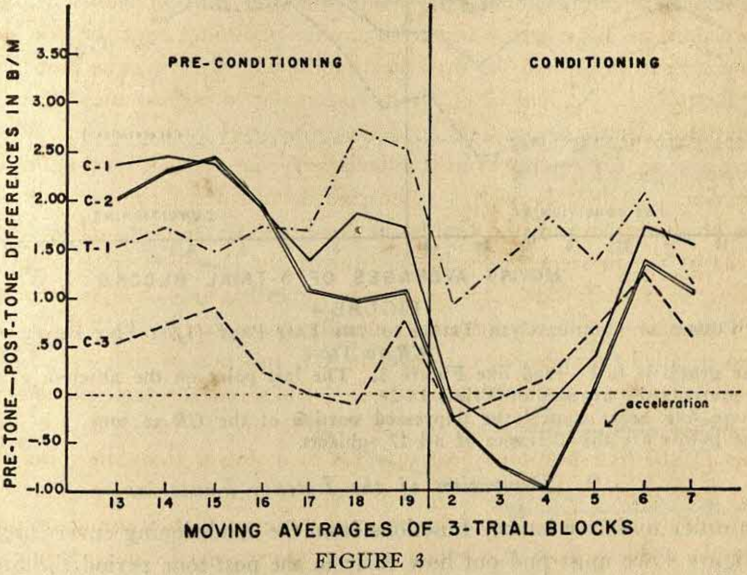


FIGURE 3

THE COURSE OF CHANGE OVER TRIALS OF THE FIRST PART OF THE CONDITIONED RESPONSE TO TONE

The ordinate expresses the difference between the value of each beat in beats per minute and the average of the last 10 beats before tone onset also in beats per minute. The abscissa is a moving average of three-trial blocks. Each block is numbered for the middlemost trial in the block. The left-hand side of the graph indicates the changes in rate for the last nine trials before conditioning trials start, while the right-hand side shows the progressive changes brought about by the first eight trials of conditioning.

The beat during tone (T-1) and the first three beats immediately after the tone (C-1, C-2, C-3) are each plotted separately. These beats control the initial, elevated portion of the *CR* to tone.

The points represent the averages of all 17 subjects.

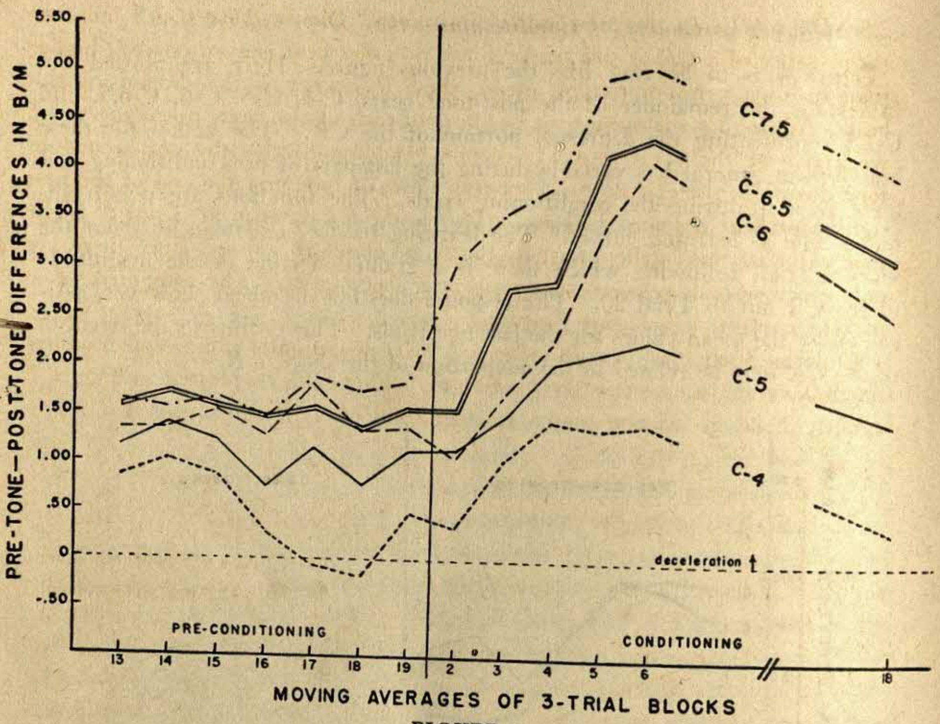


FIGURE 4
THE COURSE OF CHANGE OVER TRIALS OF THE LAST PART (LAST FIVE BEATS) OF THE *CR* TO TONE

The graph is to be read like Figure 3. The last point on the abscissa numbered 18 represents the average of Trials 16-20.

These five beats control the depressed portion of the *CR* to tone.

The points are the averages of all 17 subjects.

### 9. *Derivation of the Latency Functions*

In order to derive latency functions from the conditioning curves presented in Figure 4 we must find out how early in the post-tone period the first significant depression occurs. If we know on what beat the first significant fall from pre-tone level occurs, the ordinal beat number can be converted to a latency measure by adding up the amount of time consumed by the earlier beats. A specific example will help exposition. Consider the second block of conditioning trials in Figure 4. Which beat represents a conditioned response? The answer is that beat which is significantly different from the pre-conditioning level. The only beat which fits that description is C-7.5. When we compute the mean and *SD* of C-7.5 for the last nine pre-conditioning trials, we find that on Trial 2 of conditioning, Beat C-7.5 is more than

2.5 *SD* units from the pre-conditioning mean. Hence, beat C-7.5 indicates the presence of a *CR* at this trial. Since no other beat is even one *SD* higher than its own pre-conditioning mean, the *CR* latency is the latency of this beat. Adding up the mean time per beat of each of the beats preceding C-7.5 we arrive at the value of 6.13 seconds. Adding to this the one second of tone duration, the elapsed time between onset of tone and the end of the first significant beat becomes 7.13 sec. This is the *CR* latency at Trial 2. The true value of this latency must of course lie closer to 7.0 sec. since there are only seven seconds between tone onset and shock onset. The computed value of 7.13 is plotted in Figure 5 against Trial 2.

On trial 3, the first beat showing a 2.5 *SD* deviation is C-6.5, which occurs on the average 6.29 sec. after tone onset. This value is plotted in Figure 5 against Trial 3. If one adopts a lower level of significance, say that repre-

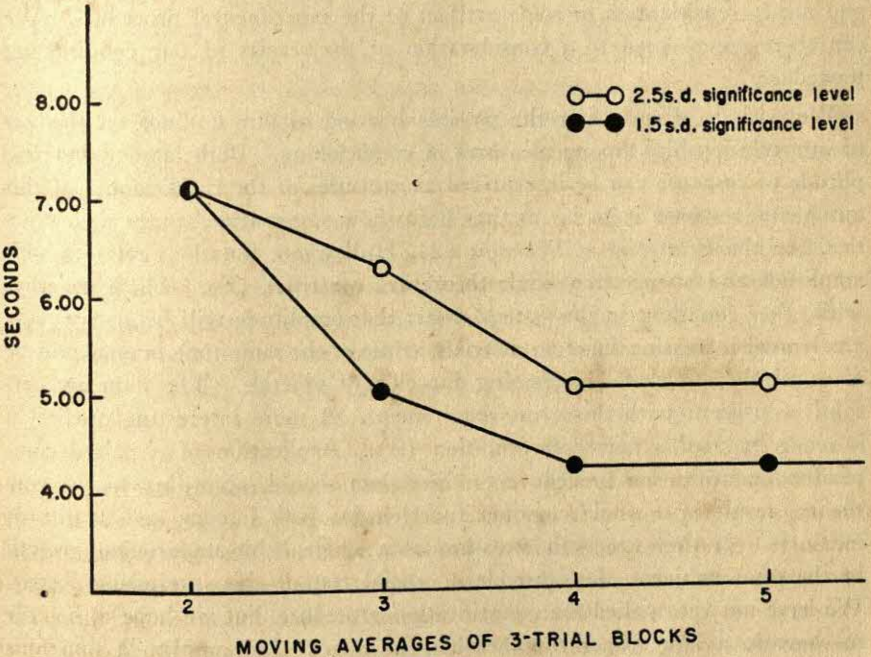


FIGURE 5

The Latency of the Conditioned Depression in Heart Rate for the First Six Trials of Conditioning

The latency is measured by the time required for the earliest beat following the tone to achieve a rate level significantly different from the pre-tone level. The two curves are derived from the use of two different levels of significance of deviation from pre-tone level.

The points are averages of 17 subjects.

sented by 1.5 *SD* unit deviations, then C-5 achieves significance on this trial. Beat C-5 is completed on the average in 5.04 sec. after tone onset. This value is plotted against Trial 3. Two functions are thereby generated, one for each of the two confidence levels employed. More curves could be computed in this general fashion, but none of them would have a latency asymptote lower than that already plotted. Any lower latency would necessarily involve Beat C-3 which does not take part in the depression effect. The over-all shape of the latency curves is one of negative acceleration with minimum points at 5 sec. (beat C-5) and 4.2 sec. (beat C-4) for the high and low significance levels respectively.

## D. Discussion

Since the results of the pseudoconditioning control test were negative, we may conclude that the observed *CR* is due to the conditioning procedure and not to sensitization or some artifact of the experimental procedure. We can therefore proceed to a consideration of the results of our conditioning procedure.

We wish to discuss first the possible bearing of our findings on theories of interrelationships among measures of conditioning. Both latency and amplitude of response can be interpreted as measures of the conditioning of this autonomic response in so far as they both show systematic changes with practice. In the *Principles of Behavior* (2), Hull wrote equations relating both amplitude and latency to a single theoretical construct, $_SE_R$. These, together with other equations in the system, assert that amplitude will be a negatively accelerated increasing function of trials, while at the same time latency will be a negatively accelerated decreasing function of practice. The data are certainly consistent with this gross requirement. A more interesting prediction is made by Hull's newer formulation (3). Application of a paired comparison quantification procedure to these data should, according to the new theory, result in a single common function for both latency and amplitude measures. If they are both measures of a single habit undergoing growth in the same subjects, they provide a crucial test of this interesting theory. We have not yet applied the quantification procedure, but we hope to be able to show in a later paper that because the latency and amplitude functions have different rates of approach to their asymptotes, the quantification procedure cannot render them identical.

Some speculations may be brought forward on the relationship of our results to anxiety. During the six seconds between tone and shock, our subjects are said by one operational definition to be in a condition of anxiety.

Schoenfeld (5) has defined anxiety as ". . . the relationship between the presentation of $S_1'$, after the $S_1$-$S_2$ operation has been sufficiently performed, and certain measurable behavior changes" (5, p. 74), where $S_1$ is a neutral stimulus and $S_2$ a noxious one. Following this notion, a possible conclusion might be that a drop in heart rate is one indicator or autonomic manifestation of anxiety. Such a conclusion must be tentative, however, until it has been shown to be true in general, that is, true regardless of the choice of specific $S_1$ and $S_2$. Consideration of two aspects of our data provides good grounds for questioning the generality of the anxiety-heart deceleration relationship. The first aspect is the secondary depression of the tone *UR*—the fact that the heart is decelerating slightly 6 seconds after the tone, regardless of the presence of shock. The effect of tone-shock pairing is to accentuate this dip. It may very well be the case that an $S_1$ other than tone which did not have an unlearned depressant effect might not result in deceleration during anxiety. A second feature of Figure 1 worth noting is the sharply dropping heart rate at the time the shock goes off. A theory will be discussed which relates this fact to the fall in heart rate after the tone without considering anxiety as at all relevant. Choosing a different $S_2$ without this feature might also yield data showing a different relation of anxiety and heart rate. We have experiments in progress testing these possibilities.

Perhaps the most interesting implications of our data are for theories of the mechanism of conditioning. Consider first the Stimulus Substitution view. According to it the conditions necessary and sufficient for conditioning are the temporal contiguity of a neutral stimulus and an established *S-R* reflex. The mechanism is simply the substitution of a neutral stimulus for the other stimulus in eliciting *R*. In the present experiment we should expect the tone to elicit heart rate *acceleration*. In view of earlier findings (4) of a conditioned *deceleration* we were surprised to find that in fact there were some aspects of acceleration in the form of the *CR*. The "initial rise effect" is a portion of our results consistent with the expectation of a Stimulus Substitution view, but it represents a weak and unstable portion of the *CR*. The major conditioning effect is the strong deceleration just before the shock, and we must look to another theory to give a clear account of this.

There are at least two general theories that will yield predictions of the obtained result. One of these is the *S-R* Contiguity view of Guthrie, the other a Hullian Reinforcement position.

Guthrie's Postremity Principle states in effect that the *last* response in a situation will recur when the situation recurs (1). One rôle of the *US* in conditioning is to radically change the situation, and thereby to protect the last response made from being displaced by an incompatible response. In the present experiment it is the shock onset which acts as the big change, protecting and thus fixating any response just prior to it in the post-tone period. Such a response would be the secondary depression of the tone *UR*. Attention to the pre-conditioning curve of Figure 1 shows that as a result of the tone, the heart is undergoing a decelerative tendency immediately before shock onset. This postreme response will be fixed during conditioning, according to the principles outlined, but the gradual changes in amplitude and latency require explanation. If it is assumed that the population of stimuli characteristic of the post-tone period contain both neutral stimuli and heart decelerating stimuli at the outset of conditioning, then the tone *UR* will appear on only a fraction of the trials (see the third postulate of Voeks, 7), depending upon the proportion of the two kinds of stimuli. The average tone *UR* for all subjects at this stage will be small in magnitude because it contains instances of no response. As conditioning proceeds, the neutral stimuli become connected to the postreme tone *UR* whenever it occurs by virtue of their contiguity with it. The result is a higher frequency of deceleration responses during the post-tone period. The effect of this on the average curve is to increase the apparent amplitude of the deceleration response during conditioning. The latency changes mean that neutral stimuli earlier in the post-tone period are being brought into contiguity with the postreme deceleration and conditioned to it.

An easy check on the validity of this type of theoretical formulation would be provided by a single change in the experimental procedure. Shortening the tone-shock interval to about 2 sec. would cause an acceleration rather than deceleration to become the postreme part of the tone *UR*. The expected outcome therefore would be an accentuation of the acceleration.

Let us look now at the relationship of our results to Hullian Reinforcement theory (2). Here the necessary and sufficient conditions for learning are given by the sequence stimulus-response-reinforcement. One rôle of the *US* is to provide the source of reinforcement. In the present experiment the termination of shock is the point of maximum reinforcement. Prior to the occurrence of shock-off, a chain of stimuli—tone and its traces, and responses—both accelerative and decelerative, can be said to exist. Hullian theory says that the response having the shortest delay of reinforcement is strengthened the most. Inspection of Figure 1 shows clearly a strong deceleration at the point of shock-off. All of the stimuli preceding this response tend to become connected to it, those closest in time to the response forming the strongest bonds. The theory would therefore predict a progressively stronger deceleration beginning with the tone and extending up to the point of reinforcement. This conditioned tendency towards deceleration is masked by the very strong acceleration of the *UR* during shock. The conditioned deceleration effect therefore can manifest itself only during the post-tone period. This prediction is in agreement with the facts. Since in Hull's theory habit acquisition is a gradual process, the observed gradual changes in amplitude and latency would also be expected.

A crucial test of this formulation can also be made by a simple change in procedure. Shortening the shock duration to about 2 sec. would cause the reinforced response to be an acceleration. If this view is correct the expected result would be a conditioned acceleration rather than deceleration.

The last explanation to be considered may be called the Anxiety *UR* hypothesis. This states simply that the temporal sequence of a neutral and an aversive stimulus creates a state of anxiety, and that heart deceleration is an unconditioned response to the state of anxiety. The results of the present experiment are of course consistent with this view. Further experimental tests of this hypothesis would consist of the demonstration of a heart rate deceleration during the *CS-US* interval regardless of the form of unlearned response to either the *CS* or aversive *US*. In the tests proposed above for the Guthrie and Hull theories, the prediction of the present view would be that no change would be observed in the form of the *CR*.

Further experimental tests of the kinds mentioned above are necessary to confirm or reject the various theories.

## E. SUMMARY AND CONCLUSIONS

Latency and amplitude characteristics of the conditioned heart rate response have been obtained from a sample of 17 subjects following a procedure introduced by Notterman, Schoenfeld, and Bersh (4). Eleven subjects were also run under a rigorous pseudoconditioning control condition. The following results were obtained.

1. A small wave-like fluctuation in heart rate was found as an unconditioned response to the *CS,* a 1-sec. tone. This response did not adapt out throughout the course of the experimental session.

2. The unconditioned response to the *US,* a mild 6-sec. shock, took the form of a rapid and strong acceleration of heart rate with partial recovery during the shock.

3. The *CR* had two aspects: (*a*) an initial acceleration, small in magnitude and unstable, and (*b*) a subsequent deceleration, larger in magnitude and stable.

4. The amplitude of the conditioned slow-down increased with negative acceleration over the first seven conditioning trials before showing a tendency toward adaptation.

5. The latencies of the conditioned drop in heart rate fell with negative acceleration to an asymptote of approximately four seconds.

6. There was little or no evidence of pseudoconditioning. The results were related to theories of interrelationships among measures of conditioning and to theories of the mechanism of conditioning.

## REFERENCES

1. GUTHRIE, E. R. Conditioning: A theory of learning in terms of stimulus, response, and association. *Forty-first Yearbook of the National Society for the Study of Education.* Bloomington, Ill.: Public School Publishing, 1942.
2. HULL, C. L. Principles of Behavior. New York: D. Appleton-Century, 1943.
3. HULL, C. L., FELSINGER, J. M., GLADSTONE, A. I., & YAMAGUCHI, H. G. A proposed quantification of habit strength. *Psychol. Rev.,* 1947, **54**, 237-254.
4. NOTTERMAN, J. M., SCHOENFELD, W. N., & BERSH, P J. Conditioned heart rate response in human beings during experimental anxiety. *J. Comp. & Physiol. Psychol.,* 1952, **45**, 1-8.
5. SCHOENFELD, W. N. An experimental approach to anxiety, escape, and avoidance behavior. In P. H. Hoch and J. Zubin (*Eds.*), *Anxiety.* New York: Gruen & Stratton, 1950.
6. SKINNER, B. F. The Behavior of Organisms. New York: Appleton-Century, 1938.
7. VOEKS, W. V. Formalization and clarification of a theory of learning. *J. of Psychol.,* 1950, **30**, 341-362.
8. ZEAMAN, D. An application of $_{S}E_{R}$ quantification procedure. *Psychol. Rev.,* 1949, **56**, 341-350.

*Department of Psychology*
*University of Connecticut*
*Storrs, Connecticut*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 251-264.

# NUTRITIONAL REPLACEMENT FOR MENTAL ILLNESS*

*Departments of Philosophy and Psychology, University of Southern California and the University of California at Los Angeles*

GEORGE WATSON AND ANDREW L. COMREY

## A. THE PROBLEM

The widespread belief in the psychogenesis of mental illness is often accompanied by a companion acceptance of the view that such illness is "mental," and consequently is not subject to physical treatment. This theory of psychic causation suggests that mental illness results from preceding mental causes and that relief from such illness would require the application of other mental causes, such as verbal interaction with a psychotherapist.

The psychogenetic theory of mental illness has achieved its general acceptance for two principal reasons. First, many cures have been effected through psychotherapy, and secondly, little success has been achieved biochemically or surgically.[1] Considerable evidence is available which conflicts with the view of psychic causation as stated. For example, it is well known that states psychologically diagnosed as neuroses and psychoses can be induced physically, e.g., in pellagra psychosis and general starvation (1, 3, 5, 6).

The experiment to be reported represents a pilot study on the effect of biochemical treatment in the form of nutritional replacement on a group of emotionally disturbed subjects. The terms "nutritional replacement" are here used to signify the systematic administration of *therapeutic* quantities of *all* the currently known or suspected vitamins and minerals, and in certain cases, the essential amino acids, needed by the human body.

None of the subjects in this experiment was mentally ill as the result of insufficient caloric intake. All were eating, and had always eaten (according to their protocols) what is commonly accepted as an adequate diet. Nutritional replacement as an hypothesis is not concerned with caloric increase in daily food intake. It is rather concerned with the differential quantitative need for certain biochemical substances.

---

*Received in the Editorial Office on June 15, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] While psychosurgery may alleviate certain symptoms, it is by no means a cure in the sense of restoring a person to normal health.

## B. PROCEDURE

Minnesota Multiphasic Personality Inventories (4) were administered to approximately 45 volunteer ambulatory unhospitalized subjects who described themselves as emotionally disturbed individuals. Thirty-four subjects were selected on the basis of *MMPI* records with one or more elevated scores. Eleven of these were assigned to the control group receiving placebos, while the remainder represented the experimental group. The assignment was at random except that five severely disturbed persons were arbitrarily assigned to the experimental group because of the gravity of their condition.[2]

The treatment for the experimental group consisted of the daily oral administration of the substances listed in Table 1.[3] These materials were made up in four capsules and four tablets, one-half to be taken after breakfast and one-half after dinner. Materials of identical appearance but with no nutritional value were supplied to the control group. Sufficient materials were given each subject to last for a month at a time. Some subjects in the experimental group received 1½ times the basic dosage per day if no effects were noticed on the regular regimen.

All subjects were interviewed at the beginning of the experiment to obtain case history information and to explain the experimental procedure. Their coöperation was sought for their own possible improvement and the potential aid to others. Each subject was given a retest on the *MMPI* approximately once each month. Some subjects began the experiment later than others and some subjects were more coöperative about completing their tests, so the number of records per subject varies. All records are included, however. For each *MMPI* record, a "total score" was computed by summing the following *MMPI* raw scores: Hs plus .5K, D, Hy, Pd plus .4K, Pa, Pt plus 1K, Sc plus 1K, and Ma plus .2K.

In addition to the test score check on the subjects' progress, signs of clinical changes were noted. The experimenters interviewed subjects briefly at each *MMPI* retest and talked with them by telephone from one to four times monthly. Reports were obtained on some subjects from members of

---

[2] All of the subjects in this experiment were either under medical care or had recently been under such care, according to their own reports. The coöperation of the doctors of individual subjects was sought in all cases.

[3] This formula was developed by W. C. Kalash, president of Kalash Vitamin Corporation, Pasadena, California. This is a *clinically* developed formula; no nutritional claims are tendered for each individual substance included, such as lithium citrate, alfalfa, kaolin, etc. The authors wish to acknowledge their indebtedness to the Kalash Laboratories for providing the materials used in the experiment.

### TABLE 1

DAILY NUTRITIONAL SUPPLEMENT

| | |
|---|---|
| Vitamin A | 30,000 U.S.P. Units |
| Vitamin D | 5,000 U.S.P. Units |
| Vitamin B-12 | 6 mcg. |
| Vitamin B-1 | 25 mg. |
| Vitamin B-2 | 20 mg. |
| Vitamin B-6 | 10 mg. |
| Vitamin C | 200 mg. |
| Niacinamide | 150 mg. |
| Niacin | 5 mg. |
| Calcium Pantothenate | 25 mg. |
| Folic Acid | 1.5 mg. |
| Rutin | 10 mg. |
| Mixed Fatty Acids (F) | 20 mg. |
| Vitamin E (Type 6-25) | 10 I.U. |
| Biotin | 25 mcg. |
| Inositol | 50 mg. |
| Choline Bitartrate | 100 mg. |
| Para-Amino Benzoic Acid | 25 mg. |
| Vitamin K (Menadione) | 0.2 mg. |
| Betaine | 25 mcg. |
| Glutamic Acid | 50 mg. |
| Lemon Peel Infusion | 15 mg. |
| Lecithin (Soya) | 25 mg. |
| Chlorophyll (100% pure) | 10 mg. |
| Methionine | 100 mg. |
| Wheat Germ | 100 mg. |
| Calcium (Dical. Phos. Lactate Anhydrous) | 1145 mg. |
| Phosphorous | 810 mg. |
| Iron (Ferrous Sulfate Exsic.) | 20 mg. |
| Iodine (Kelp) | 0.6 mg. |
| Copper (Peptonate) | 0.3 mg. |
| Manganese (Sulfate) | 10 mg. |
| Cobalt (Gluconate) | 0.5 mg. |
| Nickel (Sulfate) | 0.2 mg. |
| Magnesium (Sulfate) | 10 mg. |
| Zinc (Sulfate) | 0.2 mg. |
| Molybdenum (Trioxide) | 0.2 mg. |
| Fluorine (Bone Meal) | 0.5 mg. |
| Sulphur (Precipitated) | 1 mg. |
| Sodium Chloride | 5 mg. |
| Potassium Sulfate | 1.5 mg. |
| Lithium Citrate | 0.2 mg. |
| Pectin | 32 mg. |
| Red Bone Marrow | 25 mg. |
| Liver N.F. | 64 mg. |
| Yeast (Primary) | 64 mg. |
| Alfalfa | 100 mg. |
| Parsley | 32 mg. |
| Kelp | 32 mg. |
| Escarole | 32 mg. |
| Watercress | 32 mg. |
| Kaolin | 100 mg. |

their families. In noting clinical status and changes therein, the experimenters particularly looked for changes in emotional stability, mental confusion, orientation to the future, appraisal of self, attention to physical appearance, capacity for work, and specific symptomatology reported in the initial interview.

When it became clear that a placebo subject was showing no signs of progress after at least a month's time, he was transferred to the experimental group. This procedure represented a modification of the original experimental plan.[4]

## C. Results

Results for the placebo group are shown in Table 2. *MMPI* scores are given for each successive test period for each subject. Subjects 7, 8, 9, 14, 19, 23, and 27 showed no signs of improvement clinically after the first month and were transferred to the experimental group. Subject 1 was kept on placebos through an error for the second month, but was transferred as unimproved after that period. Subject 5 showed a slight test score improvement the first month but was transferred as unimproved at the end of the second month. Subject 22 showed an immediate test score improvement on placebos. She remained at almost the same level until the third month at which time she reported feeling worse. At the end of the fourth test period, she felt very bad and her husband seemed quite alarmed about her condition. Since her *MMPI* scores had begun to get worse too, she was transferred to the experimental group. Subject 33 showed a considerable improvement the first month, and slight gains thereafter. At the end of the last test period, she felt the treatment was "just wonderful." This subject began a reducing diet at the same time she began taking the placebos. She said that her doctor had long been advising her to reduce, but that she hesitated to do so without some nutritional support.

---

[4] This procedure was adopted for several reasons: (*a*) Many subjects indicated frank scepticism toward the proposed treatment of the experiment. Consequently, the researchers found it difficult to maintain non-suggestible subjects in the control group for even the initial period of one month. This difficulty was also experienced in the experimental group in those subjects who failed to respond to the treatment. Two subjects were lost for this reason. (*b*) The experimenters hypothesized that subjects would respond more quickly to suggestion than they would to nutritional replacement, so that controls who showed no suggestion at the end of one month were not likely to do so over a longer period. By switching such controls to the experimental group, we were able to have these subjects act as their own controls. (*c*) Since this research represents a pilot study, rather than an ideally controlled experiment, the researchers felt justified in exercising some discretion in order to maintain as large a sample as possible. Readers are reminded that this is a clinical study performed with unhospitalized subjects, rather than an animal experiment carried out under laboratory conditions.

TABLE 2

*MMPI* SCORES FOR THE PLACEBO GROUP

| Subject | Age | Sex | F | K | Hs | D | Hy | Pd | Pa | Pt | Sc | Ma | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 43 | F | 13 | 11 | 24 | 33 | 41 | 33 | 10 | 28 | 33 | 26 | 228 |
| | | | 14 | 13 | 25 | 32 | 38 | 30 | 14 | 44 | 54 | 14 | 251 |
| | | | 12 | 11 | 22 | 30 | 35 | 25 | 17 | 39 | 42 | 22 | 232 |
| 5 | 27 | F | 18 | 5 | 32 | 48 | 39 | 34 | 21 | 46 | 60 | 21 | 301 |
| | | | 26 | 3 | 31 | 41 | 35 | 29 | 19 | 48 | 65 | 26 | 294 |
| | | | 21 | 4 | 32 | 43 | 36 | 33 | 16 | 50 | 68 | 24 | 302 |
| 7 | 45 | F | 6 | 13 | 18 | 37 | 28 | 23 | 11 | 36 | 29 | 15 | 197 |
| | | | 2 | 18 | 15 | 37 | 31 | 27 | 12 | 40 | 35 | 18 | 215 |
| 8 | 40 | M | 9 | 9 | 23 | 33 | 35 | 23 | 19 | 38 | 40 | 17 | 228 |
| | | | 9 | 14 | 24 | 38 | 37 | 29 | 21 | 48 | 49 | 18 | 264 |
| 9 | 35 | F | 6 | 15 | 24 | 35 | 41 | 27 | 13 | 43 | 36 | 13 | 232 |
| | | | 2 | 17 | 21 | 36 | 41 | 29 | 10 | 38 | 33 | 12 | 220 |
| 14 | 34 | F | 5 | 11 | 13 | 26 | 27 | 29 | 14 | 44 | 41 | 24 | 218 |
| | | | 4 | 19 | 19 | 30 | 38 | 29 | 15 | 44 | 43 | 21 | 239 |
| 19 | 45 | M | 15 | 8 | 19 | 29 | 21 | 27 | 11 | 32 | 37 | 25 | 201 |
| | | | 12 | 7 | 20 | 28 | 21 | 25 | 12 | 35 | 36 | 25 | 202 |
| 22 | 27 | F | 16 | 10 | 26 | 40 | 33 | 25 | 14 | 37 | 43 | 15 | 233 |
| | | | 13 | 8 | 17 | 33 | 25 | 21 | 8 | 27 | 21 | 15 | 167 |
| | | | 7 | 10 | 18 | 28 | 23 | 20 | 6 | 23 | 18 | 13 | 149 |
| | | | 15 | 12 | 22 | 38 | 27 | 25 | 8 | 31 | 31 | 18 | 200 |
| | | | 12 | 13 | 21 | 37 | 27 | 25 | 10 | 34 | 25 | 16 | 195 |
| 23 | 35 | M | 9 | 13 | 21 | 34 | 26 | 21 | 15 | 32 | 33 | 19 | 201 |
| | | | 7 | 10 | 14 | 31 | 18 | 22 | 11 | 35 | 28 | 18 | 177 |
| 27 | 61 | M | 4 | 9 | 19 | 31 | 27 | 20 | 9 | 39 | 27 | 23 | 195 |
| | | | 9 | 8 | 20 | 34 | 24 | 25 | 10 | 38 | 28 | 27 | 206 |
| 33 | 46 | F | 2 | 22 | 21 | 32 | 35 | 30 | 14 | 37 | 26 | 16 | 211 |
| | | | 1 | 21 | 15 | 25 | 30 | 24 | 13 | 26 | 24 | 16 | 173 |
| | | | 1 | 21 | 13 | 22 | 27 | 24 | 14 | 22 | 23 | 15 | 160 |
| | | | 1 | 23 | 16 | 19 | 25 | 22 | 11 | 26 | 26 | 16 | 161 |
| | | | 0 | 22 | 15 | 21 | 27 | 24 | 11 | 26 | 25 | 16 | 165 |

Of the 11 placebo subjects, only one (subject 33) showed a sustained clinical improvement. Four showed lower values on the *MMPI* total score and seven showed higher scores, taking the difference between the first and last total score values on this treatment. The median net change in total score was an increase of one point. An increase in total score, of course, represents a more elevated, or disturbed, score. A "gain," therefore, would be represented by a *reduction* in the total test score.

Test results for the experimental group are presented in Table 3. Omitting Subject 34, who went to a mental hospital before starting the experimental treatment, there were 25 net score gains and seven net score losses. Applying the sign test (2) to these data, we determine that such a disproportionate number of improvements is significant at the one per cent level of confidence.

TABLE 3
*MMPI* SCORES FOR THE EXPERIMENTAL GROUP

| Subject | Age | Sex | F | K | Hs | D | Hy | Pd | Pa | Pt | Sc | Ma | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 43 | F | 12 | 11 | 22 | 30 | 35 | 25 | 17 | 39 | 42 | 22 | 232 |
| | | | 8 | 9 | 17 | 31 | 30 | 25 | 14 | 40 | 43 | 23 | 223 |
| | | | 9 | 10 | 12 | 29 | 27 | 21 | 11 | 35 | 33 | 18 | 186 |
| | | | 6 | 11 | 13 | 28 | 25 | 21 | 9 | 29 | 26 | 14 | 165 |
| 2 | 36 | M | 8 | 11 | 16 | 28 | 28 | 33 | 16 | 40 | 43 | 25 | 229 |
| | | | 5 | 8 | 6 | 26 | 19 | 28 | 15 | 34 | 30 | 21 | 179 |
| | | | 4 | 14 | 7 | 23 | 18 | 27 | 13 | 26 | 22 | 21 | 157 |
| | | | 2 | 22 | 11 | 16 | 22 | 26 | 11 | 29 | 32 | 24 | 171 |
| 3 | 18 | M | 8 | 23 | 15 | 30 | 22 | 25 | 12 | 38 | 39 | 14 | 195 |
| | | | 8 | 17 | 14 | 28 | 20 | 26 | 12 | 38 | 34 | 14 | 186 |
| | | | 11 | 15 | 11 | 22 | 15 | 19 | 10 | 38 | 37 | 14 | 166 |
| | | | 11 | 10 | 8 | 19 | 15 | 15 | 9 | 30 | 28 | 14 | 138 |
| 4 | 52 | M | 7 | 5 | 12 | 30 | 20 | 24 | 12 | 32 | 20 | 16 | 166 |
| | | | 5 | 10 | 10 | 23 | 20 | 27 | 6 | 23 | 19 | 19 | 147 |
| | | | 4 | 9 | 7 | 19 | 12 | 25 | 6 | 21 | 14 | 19 | 123 |
| | | | 4 | 10 | 9 | 22 | 15 | 26 | 7 | 24 | 19 | 17 | 139 |
| | | | 3 | 7 | 6 | 20 | 13 | 23 | 6 | 19 | 12 | 15 | 114 |
| 5 | 27 | F | 21 | 4 | 32 | 43 | 36 | 33 | 16 | 50 | 68 | 24 | 302 |
| | | | 21 | 3 | 32 | 38 | 36 | 30 | 17 | 48 | 61 | 25 | 287 |
| | | | 16 | 4 | 27 | 35 | 29 | 29 | 16 | 40 | 53 | 25 | 254 |
| | | | 20 | 4 | 24 | 35 | 28 | 31 | 12 | 43 | 55 | 24 | 252 |
| 6 | 75 | F | 6 | 13 | 23 | 37 | 29 | 24 | 14 | 30 | 27 | 17 | 201 |
| | | | 5 | 14 | 23 | 35 | 29 | 18 | 8 | 31 | 23 | 16 | 183 |
| | | | 4 | 11 | 17 | 31 | 21 | 16 | 6 | 23 | 18 | 14 | 146 |
| | | | 3 | 9 | 13 | 31 | 22 | 14 | 4 | 24 | 15 | 13 | 136 |
| | | | 3 | 15 | 18 | 33 | 26 | 18 | 11 | 29 | 23 | 18 | 176 |
| | | | 2 | 10 | 16 | 31 | 25 | 16 | 6 | 25 | 17 | 17 | 153 |
| 7 | 45 | F | 2 | 18 | 15 | 37 | 31 | 27 | 12 | 40 | 35 | 18 | 215 |
| | | | 6 | 14 | 15 | 38 | 30 | 27 | 12 | 34 | 31 | 16 | 203 |
| | | | 3 | 22 | 17 | 24 | 26 | 22 | 11 | 26 | 28 | 15 | 169 |
| 8 | 40 | M | 9 | 14 | 24 | 38 | 37 | 29 | 21 | 48 | 49 | 18 | 264 |
| | | | 13 | 9 | 21 | 38 | 31 | 29 | 21 | 45 | 43 | 17 | 245 |
| | | | 13 | 7 | 23 | 35 | 33 | 23 | 18 | 42 | 40 | 19 | 233 |
| | | | 13 | 9 | 26 | 34 | 34 | 26 | 15 | 43 | 43 | 19 | 240 |
| | | | 13 | 8 | 25 | 34 | 30 | 22 | 18 | 41 | 40 | 18 | 228 |
| | | | 13 | 6 | 20 | 29 | 29 | 25 | 19 | 42 | 38 | 16 | 218 |
| 9 | 35 | F | 2 | 17 | 21 | 36 | 41 | 29 | 10 | 38 | 33 | 12 | 220 |
| | | | 5 | 18 | 18 | 32 | 34 | 28 | 12 | 33 | 33 | 12 | 202 |
| | | | 4 | 18 | 18 | 30 | 33 | 26 | 10 | 30 | 31 | 8 | 186 |
| | | | 5 | 20 | 18 | 30 | 31 | 24 | 9 | 29 | 32 | 13 | 186 |
| 10 | 45 | M | 5 | 27 | 24 | 29 | 34 | 28 | 10 | 40 | 32 | 18 | 215 |
| | | | 2 | 24 | 16 | 26 | 26 | 25 | 10 | 26 | 30 | 21 | 180 |
| | | | 3 | 27 | 18 | 29 | 29 | 28 | 9 | 31 | 33 | 22 | 199 |
| | | | 1 | 24 | 17 | 23 | 29 | 26 | 11 | 26 | 28 | 21 | 181 |
| 11 | 38 | M | 1 | 17 | 22 | 28 | 34 | 27 | 14 | 30 | 22 | 21 | 198 |
| | | | 1 | 23 | 15 | 19 | 23 | 25 | 10 | 25 | 25 | 22 | 164 |
| 12 | 38 | F | 13 | 16 | 22 | 45 | 40 | 40 | 17 | 49 | 55 | 20 | 288 |
| | | | 15 | 19 | 24 | 47 | 41 | 41 | 21 | 55 | 55 | 21 | 305 |
| | | | 14 | 19 | 23 | 48 | 36 | 38 | 18 | 53 | 53 | 20 | 289 |
| | | | 8 | 22 | 21 | 31 | 32 | 30 | 14 | 41 | 40 | 20 | 229 |
| | | | 10 | 17 | 18 | 45 | 35 | 41 | 18 | 48 | 46 | 24 | 275 |

TABLE 3 (*continued*)

| Subject | Age | Sex | F | K | Hs | D | Hy | Pd | Pa | Pt | Sc | Ma | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 10 | 18 | 20 | 37 | 38 | 37 | 17 | 43 | 43 | 20 | 255 |
| 13 | 48 | F | 8 | 10 | 28 | 44 | 41 | 37 | 18 | 42 | 40 | 20 | 270 |
| | | | 5 | 14 | 26 | 35 | 34 | 36 | 12 | 41 | 27 | 17 | 228 |
| | | | 7 | 15 | 31 | 42 | 37 | 34 | 14 | 41 | 30 | 13 | 242 |
| 14 | 34 | F | 4 | 19 | 19 | 30 | 38 | 29 | 15 | 44 | 43 | 21 | 239 |
| | | | 7 | 14 | 13 | 36 | 27 | 23 | 14 | 42 | 40 | 14 | 209 |
| | | | 5 | 16 | 17 | 31 | 28 | 25 | 11 | 42 | 39 | 18 | 211 |
| | | | 6 | 12 | 15 | 29 | 32 | 29 | 10 | 36 | 37 | 22 | 210 |
| | | | 5 | 17 | 16 | 26 | 31 | 27 | 12 | 41 | 40 | 20 | 213 |
| 15 | 45 | F | 14 | 17 | 33 | 41 | 42 | 31 | 14 | 43 | 49 | 15 | 268 |
| | | | 14 | 15 | 30 | 44 | 41 | 29 | 18 | 52 | 54 | 16 | 284 |
| | | | 11 | 16 | 28 | 42 | 38 | 25 | 14 | 46 | 47 | 18 | 258 |
| | | | 8 | 14 | 27 | 40 | 38 | 24 | 12 | 44 | 39 | 13 | 237 |
| | | | 8 | 15 | 30 | 41 | 40 | 23 | 14 | 45 | 44 | 15 | 252 |
| | | | 8 | 14 | 29 | 41 | 38 | 23 | 13 | 44 | 40 | 14 | 242 |
| 16 | 46 | M | 7 | 18 | 20 | 33 | 31 | 31 | 12 | 37 | 34 | 18 | 216 |
| | | | 9 | 18 | 15 | 23 | 25 | 34 | 10 | 30 | 31 | 21 | 189 |
| | | | 7 | 20 | 17 | 24 | 24 | 33 | 11 | 34 | 29 | 20 | 192 |
| | | | 8 | 20 | 16 | 20 | 26 | 32 | 12 | 34 | 28 | 20 | 188 |
| | | | 8 | 18 | 15 | 28 | 26 | 30 | 8 | 32 | 30 | 20 | 189 |
| | | | 11 | 20 | 17 | 26 | 25 | 31 | 9 | 34 | 34 | 16 | 192 |
| 17 | 43 | F | 8 | 15 | 31 | 45 | 42 | 28 | 12 | 40 | 29 | 15 | 242 |
| | | | 6 | 19 | 27 | 46 | 34 | 31 | 10 | 39 | 28 | 16 | 231 |
| | | | 5 | 24 | 19 | 40 | 33 | 31 | 9 | 41 | 31 | 18 | 222 |
| | | | 6 | 21 | 15 | 36 | 30 | 28 | 11 | 37 | 26 | 14 | 197 |
| | | | 8 | 24 | 21 | 37 | 35 | 31 | 10 | 41 | 32 | 15 | 222 |
| 18 | 20 | F | 14 | 17 | 23 | 37 | 36 | 34 | 18 | 47 | 40 | 22 | 257 |
| | | | 13 | 15 | 22 | 41 | 31 | 34 | 16 | 50 | 44 | 16 | 254 |
| | | | 9 | 19 | 20 | 38 | 28 | 32 | 16 | 46 | 40 | 19 | 239 |
| 19 | 45 | M | 12 | 7 | 20 | 28 | 21 | 25 | 12 | 35 | 36 | 25 | 202 |
| | | | 10 | 8 | 13 | 26 | 18 | 25 | 9 | 32 | 40 | 24 | 187 |
| | | | 9 | 6 | 15 | 24 | 18 | 24 | 11 | 23 | 26 | 26 | 167 |
| | | | 9 | 10 | 16 | 26 | 21 | 25 | 10 | 31 | 33 | 26 | 188 |
| 20 | 50 | M | 6 | 16 | 34 | 39 | 40 | 25 | 14 | 41 | 37 | 13 | 243 |
| | | | 7 | 21 | 30 | 43 | 36 | 24 | 15 | 48 | 41 | 12 | 249 |
| | | | 6 | 22 | 34 | 40 | 37 | 27 | 11 | 49 | 38 | 13 | 249 |
| | | | 4 | 18 | 33 | 40 | 35 | 25 | 11 | 41 | 32 | 12 | 229 |
| | | | 5 | 20 | 30 | 38 | 34 | 27 | 13 | 42 | 34 | 10 | 228 |
| 21 | 23 | M | 4 | 22 | 16 | 24 | 27 | 37 | 13 | 27 | 37 | 19 | 200 |
| | | | 4 | 23 | 14 | 21 | 27 | 31 | 13 | 26 | 36 | 21 | 189 |
| | | | 3 | 23 | 15 | 21 | 28 | 31 | 14 | 27 | 32 | 18 | 186 |
| 22 | 27 | F | 12 | 13 | 21 | 37 | 27 | 25 | 10 | 34 | 25 | 16 | 195 |
| | | | 12 | 10 | 13 | 35 | 25 | 20 | 9 | 27 | 22 | 13 | 164 |
| | | | 14 | 8 | 14 | 38 | 26 | 20 | 10 | 30 | 30 | 14 | 182 |
| 23 | 35 | M | 7 | 10 | 14 | 31 | 18 | 22 | 11 | 35 | 28 | 18 | 177 |
| | | | 10 | 10 | 20 | 37 | 25 | 28 | 10 | 39 | 35 | 17 | 211 |
| | | | 8 | 9 | 13 | 32 | 18 | 25 | 7 | 31 | 25 | 16 | 167 |
| | | | 7 | 11 | 12 | 32 | 17 | 20 | 9 | 31 | 27 | 15 | 163 |
| | | | 5 | 10 | 11 | 30 | 22 | 23 | 11 | 24 | 29 | 19 | 169 |
| 24 | 33 | M | 2 | 13 | 23 | 35 | 33 | 24 | 10 | 36 | 26 | 21 | 208 |
| | | | 1 | 16 | 24 | 32 | 35 | 24 | 10 | 33 | 25 | 19 | 202 |
| | | | 1 | 17 | 16 | 29 | 27 | 26 | 12 | 29 | 20 | 19 | 178 |

TABLE 3 (*continued*)

| Subject | Age | Sex | F | K | Hs | D | Hy | Pd | Pa | Pt | Sc | Ma | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | 18 | 17 | 28 | 25 | 26 | 10 | 30 | 22 | 17 | 175 |
| | | | 1 | 18 | 19 | 28 | 31 | 29 | 11 | 37 | 25 | 21 | 201 |
| 25 | 68 | F | 10 | 7 | 21 | 33 | 28 | 20 | 16 | 45 | 46 | 21 | 230 |
| | | | 6 | 10 | 20 | 38 | 27 | 21 | 13 | 46 | 37 | 23 | 225 |
| 26 | 47 | F | 11 | 8 | 28 | 44 | 32 | 32 | 17 | 35 | 43 | 16 | 247 |
| | | | 10 | 12 | 32 | 41 | 36 | 31 | 17 | 45 | 42 | 19 | 263 |
| | | | 13 | 9 | 30 | 45 | 32 | 34 | 19 | 44 | 46 | 18 | 268 |
| | | | 10 | 13 | 30 | 41 | 34 | 35 | 17 | 46 | 51 | 20 | 274 |
| | | | 6 | 13 | 29 | 37 | 33 | 29 | 13 | 43 | 46 | 22 | 252 |
| 27 | 61 | M | 9 | 8 | 20 | 34 | 24 | 25 | 10 | 38 | 28 | 27 | 206 |
| | | | 16 | 9 | 23 | 37 | 30 | 27 | 11 | 39 | 35 | 24 | 226 |
| | | | 8 | 9 | 22 | 36 | 31 | 27 | 9 | 41 | 33 | 22 | 221 |
| | | | 8 | 11 | 20 | 38 | 29 | 26 | 9 | 41 | 36 | 22 | 221 |
| | | | 7 | 10 | 19 | 35 | 30 | 25 | 10 | 40 | 28 | 25 | 212 |
| 28 | 60 | M | 9 | 18 | 25 | 40 | 34 | 27 | 12 | 38 | 31 | 15 | 222 |
| | | | 6 | 22 | 25 | 36 | 34 | 29 | 12 | 45 | 34 | 19 | 234 |
| 29 | 72 | F | 4 | 14 | 26 | 39 | 41 | 28 | 16 | 45 | 35 | 14 | 244 |
| | | | 8 | 16 | 26 | 41 | 42 | 30 | 19 | 49 | 40 | 16 | 263 |
| | | | 4 | 16 | 17 | 45 | 42 | 25 | 17 | 45 | 34 | 15 | 240 |
| | | | 7 | 18 | 28 | 42 | 43 | 30 | 18 | 50 | 39 | 14 | 264 |
| 30 | 54 | M | 7 | 20 | 36 | 35 | 46 | 31 | 14 | 46 | 39 | 19 | 266 |
| | | | 7 | 16 | 33 | 39 | 46 | 29 | 15 | 47 | 33 | 19 | 261 |
| | | | 9 | 15 | 34 | 41 | 47 | 29 | 16 | 49 | 41 | 17 | 274 |
| | | | 11 | 17 | 33 | 42 | 47 | 30 | 16 | 49 | 45 | 18 | 280 |
| | | | 9 | 17 | 33 | 41 | 47 | 32 | 16 | 51 | 46 | 17 | 283 |
| | | | 12 | 17 | 33 | 41 | 44 | 32 | 15 | 51 | 50 | 23 | 289 |
| 31 | 40 | F | 12 | 11 | 32 | 43 | 44 | 28 | 11 | 33 | 31 | 24 | 246 |
| | | | 13 | 15 | 35 | 44 | 45 | 32 | 18 | 58 | 54 | 28 | 314 |
| 32 | 62 | M | 7 | 15 | 20 | 46 | 33 | 31 | 9 | 34 | 25 | 19 | 217 |
| | | | 4 | 19 | 35 | 50 | 46 | 34 | 11 | 45 | 37 | 20 | 278 |
| | | | 6 | 23 | 35 | 49 | 43 | 36 | 14 | 49 | 43 | 17 | 286 |
| | | | 8 | 21 | 34 | 49 | 44 | 33 | 19 | 49 | 46 | 18 | 292 |
| | | | 7 | 23 | 36 | 51 | 45 | 33 | 15 | 48 | 46 | 18 | 292 |
| | | | 11 | 21 | 35 | 51 | 44 | 35 | 13 | 48 | 46 | 17 | 289 |
| 34 | 42 | F | 35 | 6 | 35 | 46 | 39 | 39 | 24 | 53 | 72 | 27 | 335 |

The total score is the sum of all scores except *F* and *K*. The *K*-scale corrections are included in the tabled values for those scales where they apply. The first set of scores for each person was taken immediately before he began the experimental therapy. Those subjects who transferred from the placebo group have the same identification numbers in both Tables 2 and 3.

The median total score gain for the experimental group was 22 points. Comparing experimental and placebo groups with the Mann-Whitney Test (7), we find a difference in favor of the experimental group significant at the .03 level of confidence. A one-tailed test was used here because there was no expectation that the placebo could possibly represent an effective treatment. Conventional parametric statistics were considered unsuitable here because of the abnormal distributions of total score changes.

An examination of the clinical evidence shows that experimental Subjects 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 16, and 18 exhibited marked clinical improvement. Space does not permit a complete description of all these changes but brief case histories for a few of these subjects will be presented in the next section.

A lesser degree of clinical improvement was noted in Subjects 8, 14, 15, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 29, and 30. In these cases, there was no marked alteration of overall adjustment, despite specific improvements varying in degree and kind from subject to subject. For example, Subject 27 noticed an improvement in his condition about two weeks after he was shifted without his knowledge from placebos to experimentals. Within a month, he reported that he was sleeping better, experienced less depression, and now felt confident of getting a job. Since his test score failed to improve and some symptoms remained, he was classified as only slightly better.

There was no final evidence of sustained improvement in the status of Subjects 13, 28, and 31. Subject 13 made an initial gain but regressed and stopped the experiment because she believed the treatment was making her ill. Subjects 28 and 31 dropped out of the experiment at an early date.

In the placebo group only one case showed any sustained clinical improvement. In the experimental group, 13 cases showed marked clinical improvement. Even disregarding the 15 experimental cases showing some clinical improvement, the proportion of cases improved in the experimental group is significantly greater than that in the control group at the one per cent level of confidence. A one-tailed test was used here.

It is significant that of the 13 persons evaluated as markedly improved on the basis of clinical evidence alone, 11 were also in the highest 12 on *MMPI* total score improvement. Thus, judging by either test score changes, clinical changes, or by a combination of both criteria, significant improvement occurred in the experimental group, presumably as a result of nutritional replacement.

## D. DISCUSSION

The authors are of the opinion that the importance of this experiment is not adequately represented by the fact that a statistically significant improvement was obtained by employing a particular formula in nutritional replacement therapy. In addition, we would like to emphasize the considerations to follow.

Although the *MMPI* may record significant changes in many cases, it is misleading in others. For example, Subject 18, who was badly depressed and had suicidal intentions, completely recovered from her worst moods during the first month's treatment yet her total test score at the end of that period showed a reduction of only three points. In other words, from the *MMPI* record alone, one could not learn that this subject either did or did not have persistent and strong suicidal tendencies. Another subject, No. 30, showed an increase of 23 points after five months' therapy, but in that interval he voluntarily gave up the use of desoxyn and barbiturates, as well as a rather heavy reliance upon brandy and whisky. Although far from being recovered, he obviously had shown real improvement. Finally, several other cases showed steady clinical improvement, but their *MMPI* scores first became worse and then reversed to show an ultimate net improvement.

It is unreasonable to suppose that the experimental treatment here employed is the best that could be devised or that all individuals should have the same treatment. Under optimum nutritional replacement for each person, if such could be discovered, we believe much greater improvements could be expected. Some of our subjects were found to need adjustment in the formula, e.g., added $B_6$ (10-60 mg. daily) and amino acids (2½ grams per day of hydrolyzed yeast), which had to be provided on a crude clinical basis. Other subjects exhibited undesirable side effects, possibly resulting from the treatment, e.g., headaches, indigestion, and constipation. It was usually possible to relieve these symptoms by either reducing the dosage or adding a digestive aid.

It is difficult to estimate what further improvements would take place over longer periods of therapy. Some of our experimental subjects appeared normal, both clinically and on the *MMPI,* after three months. Others are still receiving therapy after six months and may continue to improve.

The improvement obtainable with subjects having long histories of emotional disturbance is not necessarily indicative of what might be done if corrective measures were undertaken sooner. Irreversible physical changes may make it impossible for advanced cases to respond, although Subject 6, age 75, considered hopeless by competent medical opinion, responded with marked improvement. Her original *MMPI* scores were not very extreme, however, even though she had serious clinical symptoms. (See the discussion of Subject 6 below).

A considerable portion of the experimental group in this study had previously undergone many types of therapy without substantial success. Consequently, our subjects in some respects represented an unfavorable lot from the point of view of hope for recovery. On the other hand, we had none of the very worst cases, since our sample was all ambulatory.

Better results could probably have been achieved in this study if the experimenters could have exercised greater control over their subjects. The following examples are cited as cases in point: (*a*) Subject 10 went on an extended business trip and stopped taking his dosage. He was gone for a month and lost much of his first month's improvement as a result. He regained most of his losses after another month's therapy. (*b*) Subject 12 had been released unimproved from a hospital after unsuccessful treatment, including *EST* and insulin therapy. She had a long record (about 25 years) of severe emotional illness (schizophrenia). At the end of the second month of treatment, her *MMPI* total score had decreased 59 points. Where at the start of treatment she was dazed, incoherent, disorganized, and suicidal, she was now alert, careful of her appearance, and facing adjustments at an entirely different psychological level. She was seeking ways of finding a suitable male companion, and making her own livelihood. She requested permission to talk to a psychiatrist once a month, to which the experimenters agreed, and she took a job in which there was considerable minor irritation. The experimenters next found it difficult to contact her, and when contact was made and retests obtained, it was found that she had lost about two-thirds of her initial test score improvement. Upon questioning, it was found that she had all but stopped taking the experimental dosage, and was now seeing the psychiatrist sometimes three times a week. She recognized that she was now worse and agreed to return to the experimental treatment. After two additional months of adhering to her dosage, her condition was better as indicated by both clinical and *MMPI* score evidence. (*c*) Subject 26 showed slow but steady signs of clinical improvement, although no test score improvement over her original record was noted. Specifically, a severe case of agoraphobia which kept her in seclusion was overcome, her anxiety was reduced, and vague "tensions" in her head disappeared. After two months, she felt so improved that she moved to a new residence and undertook to redecorate it herself. After this episode, which lasted about a month, she felt very bad indeed. In the meantime, she had stopped taking her dosage because she thought it might be giving her a headache, although she previously had reported that she had headaches whether she took the dosage or not.

In general, the experimenters found that as soon as the nutritional replacement was felt, the subjects began to compete on a more demanding level, often with immediate setbacks. The experimenters had no control of the subjects in this regard, and consequently feel that if adequate control over all phases of the therapy could be maintained, the results would be even more impressive.

Although some of the results of this experiment are stated in terms of median improvement, the real value of nutritional replacement as a treatment for mentally ill subjects cannot be adequately assessed in this manner. While almost all (29 of 32) experimental subjects showed some clinical improvement, a number of cases showed truly impressive gains, as illustrated by the following: (*a*) Subject 3 showed a 57 point reduction in total test score. He had been a victim of encephalitis in elementary school, and was unable to attend school at all after that. He required home care, was extremely withdrawn, and scored high on the *MMPI* schizophrenia scale. He is now attending college, majoring in chemistry, and is doing well. He is no longer withdrawn and clinically, he bears little resemblance to the person he was six months ago. (*b*) Subject 2 complained of extreme feelings of hostility, including fantasies, and was afraid of harming associates and members of his family. Highly suspicious, he was convinced he was being given "sugar pills" because there was no immediate reaction. At the end of the first month, however, he reported a marked change in feeling and attitude, finding that he was beginning to "like people" spontaneously. His clinical improvement was mirrored by a total test score reduction of 58 points. (*c*) Subject 5 presented a very disturbed *MMPI* profile which failed to improve on placebos. After eight weeks on nutritional replacement, most of her worst symptoms had disappeared. She now sleeps well, no longer experiences an "aching all over as if she had been beaten," no longer complains of vague aches and pains under stress, and her general irritability is greatly reduced. Her *MMPI* total score dropped 50 points. (*d*) Subject 1, whose total test score dropped 67 points, had a previous breakdown and was given shock treatment. At the beginning of the experiment, she complained of being afraid to go to sleep or to be alone. She felt "as though she were walking through space alone." After two months on placebos without improvement, she was switched to the experimental therapy. In two months, her fears had largely departed and she no longer felt as though she was "going to pieces." She no longer "panics" as she formerly did, her anxiety is largely gone, and little things no longer "get her down." Her last report was that she was "getting better all the time." (*e*) Subject 6 had been partly bedridden for two years and was suffering from severe depression, uncontrollable weeping, as well as visual hallucinations and extreme mental confusion. She responded rapidly on the experimental nutritional replacement, both clinically and on the *MMPI*. After one month's treatment, she was able to leave her bed for breakfast and began gardening with her extra energy. The depression largely disappeared, as did the hallucinations and mental confusion. A regression during therapy occurred when she caught a respiratory infection, but she gradually recovered this loss. The test score records accurately reflect these changes. This subject, age 75, had been given up as hopeless by literally dozens of doctors. Her family had spent thousands of dollars taking her to the famous hospitals and medical centers of the country.

The foregoing have been selected as illustrations of marked improvement and are fairly typical of others that might have been chosen.

In addition to the cases discussed above, one subject, No. 32, presents a difficult problem in evaluation. At his initial interview, he was extremely depressed and nervous, with constant thoughts of suicide. After five months on 1½ rations per day plus added amino acids, his wife reported that he was improving. For example, under mild alcoholic stimulation the subject now became cheerful and playful, as well as relaxed, where formerly such alcoholic stimulation made him more depressed. Although the subject admitted that such changes had occurred, he felt that he was still very ill and had made relatively little improvement. His test score changes would seem to support this point of view.

In a limited pilot study of the kind being reported, it is obvious that the total range and complexity of problems generally classed as emotional disturbances cannot be explored. The experimenters feel, however, that sufficient evidence of a positive kind has been gathered to warrant the conclusion that some types of emotional illness apparently respond markedly to nutritional replacement.

Although the research here reported is to some degree exploratory in nature, the hypothesis of nutritional replacement was evolved through previous unpublished research extending over several years. Even if the present experiment had conformed to the ideals for controlled research in every respect, the need for further tests of the hypothesis would still remain. The experimenters wish to emphasize that the hypothesis of nutritional replacement in relation to emotional illness can only be confirmed fully through further research by independent investigators. In reporting their research at a relatively early stage, the authors hope to interest others in testing this hypothesis on a wider and better controlled scale, such as under hospital conditions.

The experimenters offer no speculation as to the nature of the bodily processes involved in the changes apparently produced by the experimental treatment employed. The further problem of the relation between conventional psychotherapeutic techniques and nutritional replacement therapy is too complex to be discussed in the present report. The authors wish to suggest, however, one tentative conclusion in theoretical psychology. We feel that the experimental evidence reviewed here may justifiably be considered as partial confirmation of the hypothesis that some mental illness can be at least partially relieved by employing purely physical, as contrasted to psychological, therapy.

## REFERENCES

1. BROZEK, J., & ERICKSON, N. K. Item analysis of the psychoneurotic scales of the Minnesota Multiphasic Personality Inventory in experimental semi-starvation. *J. Consult. Psychol.*, 1948, **12**, 403-411.
2. DIXON, W. J., & MASSEY, F. J. Introduction to Statistical Analysis. New York: McGraw-Hill, 1951.
3. FRANKLIN, J. C., & BROZEK, J. The Rosenzweig P-F test as a measure of frustration response in semi-starvation. *J. Consult. Psychol.*, 1949, **13**, 293-301.
4. HATHAWAY, S. R., & MCKINLEY, J. C. Manual for the Minnesota Multiphasic Personality Inventory. Minneapolis: Univ. Minnesota Press, 1943. (Now sold by the Psychological Corporation, New York.)
5. KEYS, A., *et al.* The Biology of Human Starvation. Minneapolis: Univ. Minnesota Press, 1950.
6. KJENAAS, N. K., & BROZEK, J. Personality in experimental semistarvation. *Psychosomat. Med.*, 1952, **14**, 115-128.
7. WALKER, H., & LEV, J. Statistical Inference. New York: Holt, 1953.

*Department of Philosophy*
*University of Southern California*
*Los Angeles 7, California*

*Department of Psychology*
*University of California*
*Los Angeles 24, California*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 265-270.

# THE EFFECT OF FIGURAL REPRODUCTION ON RECOGNITION*

*San Diego State College and the U. S. Navy Electronics Laboratory, San Diego*

MERLE B. TURNER AND EUGENE A. CRAIG

## A. PROBLEM

Does the figural "trace" undergo change in the direction of "better" (regularity, simplicity, symmetry) configuration? If so, is the change autonomous and more or less inherent in the figural system itself? (6, 9). Or, is the change attributable to influences from sets, traces, associations, etc., external to the trace in question? (1, 2, 3, 4, 10). If the change is due to intrusions from other systems (1, 2, 3, 8), then the experimental paradigms employed in studies of proactive and retroactive inhibition are relevant to the problem.

Hanawalt (3) has indicated that recognition is a more valid test of mnemonic change than is reproduction, since reproduction involves motor skills apparently extrinsic to a purely perceptual trace. Others (4, 7) have accepted Hanawalt's suggestion, but in so doing, have eliminated the influence of motor-perceptual systems upon specific memory traces. The question remains, then, as to whether motor actions in the reproduction of figures do actually constitute intrusions upon the primary perceptual trace. Hanawalt's criticism loses force if such effects cannot be demonstrated. Conversely, if these effects are found, it is of special interest to explore their nature.

A reasonable hypothesis in the light of the earlier studies is that, if mnemonic change does result from motor intrusions, its direction will be toward greater regularizations of the figural trace. A further relevant experimental question is whether any such intrusions persist or whether, with the lapse of time, the interaction loses effectiveness.

The present writers believe these hypotheses may be effectively tested by employing the reproduction of figural stimuli as interpolated materials in the familiar $A_1 BA_2$ scheme, where $A_1$ represents learning, $A_2$ recognition and B represents interpolated materials.

---

*Received in the Editorial Office on June 15, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

## B. METHOD

### 1. *Subjects*

*S*s were 259 students at San Diego State College. Classes in introductory and social psychology were employed as groups for the experiment since we had no reason to suspect that ability to recognize would be a function of class differences. The classes were undergraduate and did not differ significantly in age distribution or sex distribution.

### 2. *Stimulus and Recognition Figures*

Figural designs were constructed by the superposition of two quadrilaterals differing in orientation to one another. The regularity ("goodness") of the resulting figures varied in terms of the extent of deviation from parallel structure. It was decided to employ only one of these variations as the stimulus figure in order to avoid the intra-serial effects mentioned by Gibson (2).

Two slides were made up. One contained the five figures employed in the recognition situation. Figure 1 shows these in the order in which they appeared on this recognition slide. Below each figure is a number referring to its rank on a scale of relative regularity. (These numbers, of course, did not appear on the actual recognition slide). The figure ranked 3 was reproduced on a separate slide and used as the stimulus figure. It should not, however, be regarded as a neutral figure since it is somewhat closer to the less regular (4 and 5) than to the more regular figures (1 and 2). This arrangement was chosen to insure that any bias exhibited by the recognition slide would be in a direction opposite to that of our hypothesis.

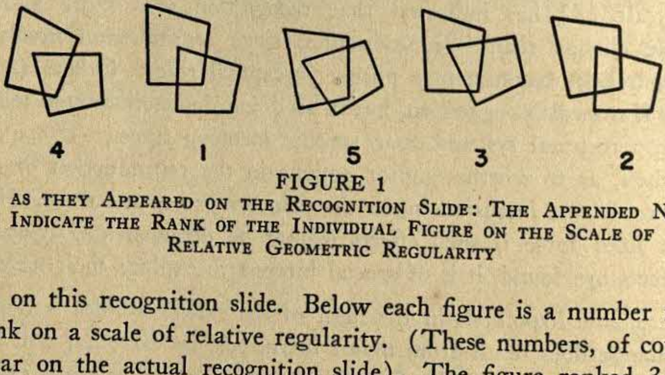


FIGURE 1
FIGURES AS THEY APPEARED ON THE RECOGNITION SLIDE: THE APPENDED NUMBERS INDICATE THE RANK OF THE INDIVIDUAL FIGURE ON THE SCALE OF RELATIVE GEOMETRIC REGULARITY

### 3. *Procedure*

The figures were presented by means of a projector equipped with a shutter for tachistoscopic exposure. Five groups of *S*s were used, three experimental and two control. *S*s in all groups received three .04 second exposures of the stimulus figure (learning trials) and subsequently received the recognition slide. *S*s task was to attempt to select the stimulus figure from among the five recognition slide figures. The following groups were used:

*a. Group 1.* *S*s in this experimental group received the learning trials then immediately reproduced the stimulus figure. They had one minute for the task, and then performed the recognition task.

*b. Group 2.* This group acted as a control for Group 1. Conditions were equivalent for both groups except that *S*s in the control group did not reproduce the stimulus figure.

*c. Group 3.* This experimental group reproduced the stimulus figure immediately after the learning trials, then performed the recognition task after a delay of 48 hours.

*d. Group 4.* This experimental group reproduced the stimulus figure 48 hours following the learning trials, then performed the recognition task.

*e. Group 5.* This group acted as a control for Groups 3 and 4. The *S*s received the recognition slide 48 hours following the learning trials, but did not reproduce the stimulus figure.

In all groups, experimental *S*s were exposed to the interpolated activity of reproduction; the control *S*s were not. Instructions to all *S*s, irrespective of group, prepared them for the task of recognition, but not for that of reproduction. No reference was made in regard to when they were to receive the recognition slide.

## C. RESULTS AND DISCUSSION

A fourfold contingency table was used to facilitate group comparisons. Raw frequency data, which yield the percentages of Table 1, were categorized into *S*s choosing more regular figure and *S*s choosing not-more regular figure

TABLE 1

PERCENTAGE OF SUBJECTS SELECTING MORE REGULAR AND NOT-MORE REGULAR FIGURE ON RECOGNITION TEST

| Group | Percentage selecting More regular | Not-more regular | N's |
|---|---|---|---|
| 1. I.Rp. — I.Rc.* | 43 | 57 | 49 |
| 2. N.Rp. — I.Rc. | 19 | 81 | 73 |
| 3. I.Rp. — D.Rc. | 31 | 69 | 45 |
| 4. D.Rp. — D.Rc. | 50 | 50 | 44 |
| 5. N.Rp. — D.Rc. | 31 | 69 | 48 |
| N's | 86 | 173 | 259 |

Note: — N's relate to marginal totals.
*I. indicates immediate; D., delayed; N., no; Rp., reproduction; Rc, recognition.

(same or less regular) than they had been exposed to during the learning trials. No differentiation was made between the two figures that are more regular than the stimulus, nor between the two less regular figures. If choice of figure on the recognition slide were a matter of chance, then the expected percentage frequency would be 40 for more regular and 60 for not-more regular.

A comparison of Groups 1 and 2 of Table 1 yields information relating to the immediate recognition situation. The fourfold categorization of the raw frequency data for these two groups yields a $t$ of 2.84 ($p = .002$).[1] This indicates that for this particular comparison the tendency to recognize a more geometrically regular figure than the original stimulus was a function of the interpolated activity of reproduction.

An examination of the three delayed recognition groups gives information relating to the effects of the temporal placement of the activity of reproduction. There is no difference between the immediate reproduction and no reproduction groups (3 and 5). However, a comparison of the delayed reproduction and no reproduction groups (4 and 5) yields a $t$ of 1.83 ($p = .035$); and a comparison of the immediate and delayed reproduction groups (3 and 4) yields a $t$ of 1.82 ($p = .035$). It would appear, then, that the interpolated activity of reproduction enhances the error in the direction of recognizing more regular figure primarily when the reproduction occurs just prior to the recognition task.

A comparison of the two control groups (2 and 5) gives a $t$ of 1.52 ($p = .06$). This suggests that errors in terms of recognizing more regular figure may increase when the period of delay between the stimulus presentation and the recognition is increased beyond one minute. If this effect is real, it is amenable to interpretation either in terms of confirming autonomous trace dynamics or in terms of hypotheses of multiple trace interaction such as have been proposed by several writers (1, 2, 4). It is difficult to see, however, how the Gestalt doctrine of autonomous trace dynamics can conveniently account for the previously mentioned effect of interpolated reproduction on recognition. It appears that the motor-perceptual act has in some manner interacted with non-isolated perceptual traces to effect the shift in recognition.

Stimulus impressions may be thought of as being recorded in more than one trace system. Verbal, motor, motivational, and affective systems all modify the nature of the trace, and hence, recognition of the figure. That *S* recognizes more regular figure than he originally sees presumably is due to the fact that the primary trace has been modified in this direction by other trace systems.

---

[1] All $t$'s are derived from the square root of chi-square. A one-tailed test is applied (5).

The present research attempts only to ascertain the influence of motor-perceptual systems upon the primary trace. It does this by interpolating reproduction at two different intervals between stimulus presentation and recognition. For the most part, it appears that reproduction has its greatest impact upon the primary perceptual trace when it occurs just prior to the act of recognition. This is in keeping with the well-known effects of recency. When the motor-perceptual trace is relatively more recent, it will exert a greater influence upon the primary perceptual trace to which it is related.

## D. Summary

This study attempts to ascertain the effects of interpolated reproduction upon recognition, when the dependent variable is error in recognizing more regular (better) figure than that to which *S* is exposed in the training trials. Five groups of *S*s were used, two participated in conditions of immediate recognition, three participated in conditions of delayed recognition. All groups received three .04 second exposures of a stimulus figure. Group 1 reproduced the figure immediately after the stimulus exposure, then was presented with the recognition task. Group 3 reproduced immediately, then performed the recognition task after a delay of 48 hours. Group 4 delayed reproduction for 48 hours, then immediately thereafter performed the recognition task. Groups 2 and 5 served as controls who performed the recognition task, but did not reproduce the stimulus figure.

1. In immediate recognition, interpolated reproduction tends to increase errors in the direction of recognizing more regular figure than that of the original stimulus.
2. In delayed recognition, interpolated reproduction tends to increase errors in recognizing more regular figure only when the act of reproduction occurs just prior to the act of recognition.

The results are interpreted in terms of the interaction between a primary perceptual and a motor-perceptual system.

## References

1. Carmichael, L., Hogan, H. P., & Walter, A. A. An experimental study of the effect of language on the reproduction of visually perceived form. *J. Exp. Psychol.*, 1932, **15**, 73-86.
2. Gibson, J. J. The reproduction of visually perceived forms. *J. Exp. Psychol.*, 1929, **12**, 1-39.
3. Hanawalt, N. G. Memory trace for figures in recall and recognition. *Arch. Psychol.*, 1937, No. 216. Pp. 89.

4. Hebb, D. O., & Foord, E. N. Errors of visual recognition and the nature of the trace. *J. Exp. Psychol.,* 1945, **35**, 335-348.

5. Jones, L. V. Tests of hypotheses: one-sided vs. two-sided alternatives. *Psychol. Bull.,* 1952, **49**, 43-46.

6. Koffka, K. Principles of Gestalt Psychology. New York: Harcourt Brace, 1935.

7. Seidenfeld, M. A. Time as a factor in the recognition of visually perceived figures. *Amer. J. Psychol.,* 1938, **51**, 64-82.

8. Woodworth, R. S. Experimental Psychology. New York: Holt, 1938.

9. Wulf, F. Über die Veränderung von Vorstellungen (Gedächtnis und Gestalt). *Psychol. Forsch.,* 1922, **1**, 333-373.

10. Zangwill, O. L. An investigation of the relationship between the processes of reproducing and recognizing simple figures, with special reference to Koffka's trace theory. *Brit. J. Psychol.,* 1937, **27**, 250-276.

*Department of Psychology*
*San Diego State College*
*San Diego 15, California*

*U. S. Navy Electronics Laboratory*
*San Diego 52, California*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 271-280.

# SELF-DISLIKE AS A DETERMINANT OF MARKED INGROUP-OUTGROUP PREFERENCES*[1]

*Department of Psychology, Boston University, and Center for International Studies, Massachusetts Institute of Technology*

ARTHUR J. BRODBECK AND HOWARD V. PERLMUTTER

## A. THE PROBLEM

Since the publication of *The Authoritarian Personality* (1), there has been a continuous research interest in the personality determinants of ingroup-outgroup attitudes and behavior. Emphasis has especially been given to the personality determinants of a special pattern of pro-ingroup and anti-outgroup attitudes: *ethnocentrism,* usually found to correlate highly with authoritarianism. Recently, a series of studies have attempted to isolate a different and opposite kind of ingroup-outgroup preference—i.e., one in which there is an anti-ingroup and pro-outgroup preference. This attitude has been labeled *xenophilic* by Perlmutter (6). In both kinds of patterns, and with the measuring devices used, there is marked use of ingroup-outgroup distinctions, but with valuations and preferences in opposite directions. There is some evidence to suggest, oddly enough, that the *same* personality dynamics lead an individual to adopt either extreme preference, at least for groups whose information about outgroups is limited. Those who possess xenophilic attitudes have been shown to share with ethnocentric individuals some of the same personality dimensions emphasized by the California group of investigators (6).

The present study is concerned with one specific, but quite general, personality variable which conceivably may be of central importance to the construction of a theory regarding *any form* of marked ingroup-outgroup preference. It is a personality variable which, furthermore, may be acquired during early socialization experiences within the family. The parent-child relationship is presumably one in which basic ideas and feelings about the self are established (8). If the socialization process has been successful, then we would expect identification with parents to lead the individual to have positive attitudes toward the self, to possess high self-esteem and self-confidence. On the other hand,

---

*Received in the Editorial Office on June 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This is one of a series of studies of "Personality Determinants of International Communications" being conducted in the International Communications Program at the Center for International Studies, Massachusetts Institute of Technology.

if the parent-child relationship results in unresolved conflict, then we would expect the conditions of socialization would be such as to favor the acquisition of negative attitudes toward the self—self-dislike—unless there are other socialization forces counterbalancing those of the family. It is possible that negative and positive attitudes toward the self is a personality dimension created by parental socialization and partially determines adult acceptance of cultural ideologies involving marked pro- and anti-ingroup and outgroup preferences. Whiting and Child (9) have suggested such a conceptual model for the analysis of personality-culture materials. Strong self-dislike would, according to the hypothesis put forward here, lead to overidentification with, or rejection of, ingroup authority; the valuation of ingroup or outgroup would be markedly one-sided.

According to such a theory, ingroup-outgroup preferences arise from any kind of conflictful parent-child relationship *per se,* but which *form* the preference takes—ethnocentric or xenophilic—depends in part upon the exact nature of the parent-child conflict. The tendency to dislike the self would presumably be a consequence of *any* kind of parent-child conflict and, hence, should be related to *any* form of marked ingroup-outgroup preference and overevaluation. It is this last hypothesis, deduced from the socialization theory, that this study attempts to investigate. *A dislike of the self,* which it is presumed arises from socialization experiences in the family, will be associated with either consistent outgroup or consistent ingroup preferences.

There is some research which already supports this hypothesis. Perlmutter (5) has given evidence that low self-esteem is significantly related to desire to live abroad. If low self-esteem is taken as an approximate measure of dislike of the self, then this finding is partial confirmation for the hypothesis to the extent that desire to live abroad indicates a preference for an outgroup. Brodbeck (3) has found that individuals with strong negative self-attitudes with respect to special traits are more prone to perceive sharp negative and positive differences between themselves and others, and less likely to find common resemblances between themselves and others, for the same traits. The result was found, however, *among* ingroup members—i.e., a fraternity group—who knew each other well, and thus, is only very partial and indirect confirmation for the hypothesis. However, the study does suggest that individuals with marked self-dislike would be more disposed to make exaggerated social discriminations.

## B. PROCEDURE

In order to test the central hypothesis, a series of 78 items, all of which were designed to tap strong feelings of self-dislike, were administered to an undergraduate evening course in child psychology at Boston University, consisting of 22 females and 14 males. Because our hypothesis concerned *self*-dislike, the items were devised so as to exclude as far as was possible the direct measurement of anxiety. Although the test logic rests upon honest self-reporting of feelings of self-dislike, self-reporting may be affected obviously by either (*a*) repression and suppression or (*b*) unwillingness to admit to another feelings one knows one has. The latter factor was presumably reduced since anonymity of the subjects was assured. Because of this, the test should be looked upon as a measure of willingness and capacity *to report feelings of self-dislike*. Other kinds of techniques will be required to estimate the manner in which the factors of repression and/or dishonest reporting enter generally into the test results.

A correlation was run of each of the 78 items with the total test score. Only those items which correlated .40 or better were retained for the final scale used in this study. The items were as follows:

1. I am seldom able to laugh at myself and my shortcomings.
2. I usually take the blame upon myself when things go wrong.
3. I regret that I have not been better to my parents in the past than I have.
4. I usually get punished for feeling too happy or too self-confident.
5. I find it hard to believe compliments given me by others.
6. I usually tend to withdraw from situations where my real feelings might be understood.
7. One becomes a better person through some amount of suffering in life.
8. I tend to be rather critical of myself, sometimes even criticizing myself to others.
9. I wish I could believe in myself more than I do.
10. I find it hard to forget the times when I have done a wrong.
11. I let good opportunities go by without taking full advantage of them for myself.
12. I get furious with myself after times when I have acted silly and like a child.
13. I still have a long way to go to become the kind of person I want to be.
14. I tend to examine my past behavior over and over again.
15. When I disagree with others, I feel deep down inside me that I am perhaps wrong, even though I may not say so or give in.
16. I feel I do not have much to be proud about.
17. I tend to doubt my own explanations as to why I act the way I do.
18. I find it hard to keep from "giving myself away" to others.
19. I feel much better if I have to pay for my mistakes.
20. It is hard for me to be at ease with important people.

21. I sometimes feel all bottled-up inside, as if part of me could not get out or be expressed.
22. I certainly know what it is like to feel unclean inside.
23. I seldom feel that I deserve the good things that come my way.
24. I am not good at pretending; people see through me and my pretense very easily.
25. I get embarrassed very easily.
26. I do not like to put my abilities to the test.

These items were administered to the 141 females at the School of Practical Arts and Letters and 132 males from the College of Business Administration, both of which also took the *XPL* test, discussed in a previous paper (6), and items from the Adorno, Frenkl-Brunswik scale for authoritarianism. The *XPL* test consisted of a series of 10 relatively dogmatic statements wherein comparison between European and American individuals and institutions were made. A high *XPL* score (on a seven-point scale) indicates consistent agreement with the statements *favoring* the European "way" over the American. The *XPL* questionnaire was as follows:

1. The chief stimulants to basic American institutions in this century have come mainly from European and Oriental ideas and doctrines.
2. The most delicious foods are European or Oriental.
3. The British use the English language better than do most Americans.
4. European men are usually more romantic than American men.
5. Too much stress is placed on Americanism nowadays, not enough on the fact that we are all descendants of the Old World.
6. Compared to the French, Americans are an unimaginative people.
7. European children are generally better mannered than Americans.
8. Europeans on the whole appreciate and understand the arts better than Americans.
9. Most European girls make better wives than American girls.
10. Europeans generally are a warmer and friendlier people than Americans.

The seven-point scale used for these latter tests were used for the Self-Dislike (*SD*) questionnaires. The Self-Dislike questionnaire was readministered and found to have a test-retest reliability of .8.

## C. Results

According to our theory, individuals with marked ethnocentric, authoritarian attitudes or xenophilic attitudes ought to be higher in feelings of self-dislike than individuals with no marked ingroup-outgroup preferences. Table 1 shows the results for high, intermediate and low xenophiles among the

TABLE 1
COMPARISON OF SELF-DISLIKE SCORES (*SD*) FOR HIGH, INTERMEDIATE, AND LOW XENOPHILES

| | *N* | Mean *XPL* score per item | Mean self-dislike score per item | Hi vs. Lo *XPL* |
|---|---|---|---|---|
| | | *CBA Males* (*N* = 132) | | |
| High *XPL* | 29 | 4.9 | 4.2 | .01 |
| Intermed. *XPL* | 69 | 3.4 | 3.8 | |
| Low *XPL* | 34 | 2.4 | 3.5 | |
| | 132 | | | |
| Average: | | 3.5 | 3.8 | |
| | | *PAL Females* (*N* = 141) | | |
| High *XPL* | 30 | 5.1 | 4.0 | .08 |
| Intermed. *XPL* | 78 | 3.7 | 3.6 | |
| Low *XPL* | 33 | 2.5 | 3.6 | |
| | 141 | | | |

*CBA* males and *PAL* females. It is clear that the scores for both men and women are in the predicted direction. The scores for the males are highly significant, whereas the scores for the females approach significance at a little less than the .08 level of confidence, using a two-tailed *t*-test. Table 2 summarizes the relationship between authoritarianism and self-dislike among both the males and females. Again, both sets of data are in the predicted

TABLE 2
COMPARISON OF SELF-DISLIKE SCORES FOR HIGH, INTERMEDIATE, AND LOW AUTHORITARIANS

| | *N* | Mean *A* score per item | Mean self-dislike score per item | Level of signif. Hi vs. Lo *A* |
|---|---|---|---|---|
| | | *CBA Males* (*N* = 132) | | |
| High *A* | 29 | 5.1 | 4.3 | .01 |
| Intermed. *A* | 70 | 3.4 | 3.7 | |
| Low *A* | 33 | 2.8 | 3.5 | |
| | 132 | | | |
| Average: | | 3.6 | 3.8 | |
| | | *PAL Females* (*N* = 141) | | |
| High *A* | 30 | 4.9 | 3.8 | .08 |
| Intermed. *A* | 71 | 4.1 | 3.6 | |
| Low *A* | 30 | 3.0 | 3.4 | |
| | 141 | | | |
| Average: | | 3.7 | 3.6 | |

direction, but the relationship of the variables among the females is less so. There may be a sex factor operating here, since it seems that, for both kinds of preference patterns, the women do not show as strong a relationship as the men do for reported feelings of self-dislike.

In order to throw further light on the differential rôle the sex factor might play in the relationship between the variables, an analysis was made of the extent to which authoritarianism or xenophilic attitudes can be predicted from the self-dislike questionnaire, rather than vice-versa. The relevant data is summarized in Table 3 for both males and females. In this analysis, the sex-factor becomes even more prominent. While the self-dislike scores differentiate the extremes of the high and low scoring authoritarian and xenophilic men, it does not do so as clearly for women. As a matter of fact, the women with middle or moderate xenophilic scores (mid *XPL*) tend to score as less prone to feeling of self-dislike than the low scoring xenophilic women (low *XPL*).

TABLE 3
COMPARISON OF AUTHORITARIAN AND XENOPHILE SCORES FOR HIGH, INTERMEDIATE, AND LOW SELF-DISLIKE SCORERS

| | *N* | Mean self-dislike score per item | Mean *A* score per item | Level of signif. | Mean *XPL* score per item | Level of signif. |
|---|---|---|---|---|---|---|
| | | *CBA Males* (*N* = 132) | | | | |
| High *SD* | 32 | 4.9 | 4.0 | .01 | 3.8 | .01 |
| Intermed. *SD* | 71 | 3.6 | 3.9 | | 3.5 | |
| Low *SD* | 29 | 3.1 | 3.3 | | 3.2 | |
| | 132 | | | | | |
| Average: | | 3.8 | 3.5 | | 3.8 | |
| | | *PAL Females* (*N* = 141) | | | | |
| High *SD* | 31 | 4.7 | 4.0 | n.d. | 3.8 | n.d. |
| Intermed. *SD* | 80 | 3.7 | 3.6 | | 3.5 | |
| Low *SD* | 30 | 2.6 | 3.8 | | 3.5 | |
| | 141 | | | | | |
| Average: | | 3.7 | 3.7 | | 3.6 | |

It is remembered that the self-dislike test is a test of *willingness* to report such feelings; repression may, therefore, affect the test scores. It is possible that self-dislike is as highly related to authoritarian scores and xenophilic scores for the women as it is for the men, but the women are more unable or more unwilling to report such feelings. There are two ways to test such an hypothesis. We would expect, if the women were more unwilling to report such feelings, that women would have lower self-dislike scores than men. The difference between the mean scores for the men and women are shown in Table 3. The differences are not significant. The result, therefore, offers no confirmation for the hypothesis. It may be, however, that unwillingness to report feeling of self-dislike are specific to authoritarian and xenophilic women rather than to women in general.

From Table 1 it is clear that there are no differences between high *XPL* men and women with respect to their scores on the self-dislike test. Furthermore there is little difference between low *XPL* men and women on scores of self-dislike. The differences seem to appear with respect to the high and low Authoritarian males and females as seen in Table 2. The high *A* males have significantly higher scores on self-dislike than the high *A* females, while the low Authoritarian males have a slightly higher self-dislike score even though the mean *A* score of the females is a little higher.

In addition, from Table 3 it can be seen that while there is no difference between the mean of the upper quartile of *SD* scores (high *SD* scores), the mean of the lower quartile *male* scores (low *SD* scores) is significantly higher than those of the *females* at the .01 level of significance. On the other hand, there are no differences between the *A* scores for either high *SD* males and females or low *SD* males and females. A difference is found between the *XPL* scores of low *SD* scorers, the males scoring significantly lower on the *XPL* scale than the females.

## D. DISCUSSION

The results of this study would tend to indicate that self-dislike is associated with authoritarian-ethnocentric and xenophilic attitudes, at least for males. For females, the results are much more ambiguous, and seem to be complicated by other factors. It has been suggested by Ausubel (2) and Parsons (4) that women may acquire a higher degree of self-dislike in the socialization process than men. This factor may operate in such a way as to raise the strength of self-dislike in xenophilic and authoritarian women to such a degree above that of males that repression and/or unwillingness to report such strong feelings are instigated. On the other hand, it may be that self-dislike is universally so high among high and low *XPL* and Authoritarian women that it is not a decisive factor in determining ingroup-outgroup tendencies as it seems to be in males. Obviously more research is needed to decide upon these alternatives. It might be that a test of self-esteem, considered the more positive end of a continuum of which self-dislike is the negative end, would show the differences between high and low *XPL* and Authoritarian females more clearly. Women, in other words, may be more easily differentiated in terms of degree of self-esteem which they possess, while men are more easily differentiated by degree of self-dislike, or are differentiated on both ends of the continuum. These two hypotheses—that women are more prone to repression of guilt and that they may be more easily differentiated from each other in terms of self-esteem—are not incompatible; they may both operate.

It has been suggested that parental influence determines the attitudes toward the self, whether they be those of self-dislike or self-esteem. Both high *XPL* and Authoritarian scorers have been conceived to have conflictful parental relationships, and unsuccessful socialization experiences within the family, thus leading to self-dislike. On the other hand, low *XPL* and Authoritarian scorers have been conceived to have relatively more satisfactory and successful socialization experiences within the family. Such a socialization hypothesis needs greater specificity in order to be tested further. It is especially important to know what are the differences in the socialization disturbances that produce, on the one hand, an ethnocentric outcome, on the other, a xenophilic outcome.

The ethnocentric outcome is conceived to arise from *direct* experience of negative evaluations, instigated by parents. The parent, so to speak, has had difficulty accepting the child, leaving him with perceptions of strong difference and distance between himself and his parents. The ethnocentric individual's strong identification with the ingroup is a way in which he attempts to rid himself of marked feelings of "psychological distance" built up in past experience between authority and himself, a way in which to assuage his negative self-attitudes picked up from the parents by becoming identified with the authoritarian source. His preference pattern is, therefore, not based upon the objective properties of the ingroup, but upon a guilt-assuaging device and, therefore, is often highly emotional and defensive in nature, as the California investigators have suggested.

The xenophilic outcome, however, may, under some conditions, arise from a different kind of conflict. It occurs as a result of not having an adequate parent to identify with; where the socializer has in some way been perceived to be inadequate and has communicated feelings of self-dislike to the child. Because of this, identification with the socializer *induces* rather than *reduces* self-dislike. Identification with the out-group offers a way of resistance, and some security for rebellion, against identification with the socializer which has itself in the past only produced further self-dislike.

These hypotheses regarding the differential antecedent socialization conditions producing the two different outcomes must be qualified in many ways and certainly qualified to include adolescent experiences. A xenophilic adolescent may shift toward ethnocentrism by being accepted into a peer-group of similar characteristics to those of his parents and by being rejected by peer groups of preferred, but dissimilar, characteristics to those of his parents. Similarly, an ethnocentric child may shift toward a xenophilic attitude by rejection by a peer group of similar characteristics to those of his parents and by ready acceptance into a peer group of quite dissimilar characteristics. No doubt extreme preferences may be behaviorally and internally modified by such complications not only in adolescence but to some degree throughout life, making for more equalitarian personalities. The degree to which childhood socialization may be modified in adolescence needs further research. Studies which are quoted by Ausubel (2) seem to indicate that these changes that may occur are greater than many personality theories take into account.

Some support for the socialization hypothesis as outlined in this discussion comes from a study by Sarnoff (7). This researcher found that Jewish students with strong self-hate were likely: (*a*) to perceive their parents as inadequate and weak models, (*b*) to have marked anti-Semitic tendencies. This finding seems especially important when contrasted with the results from the original California studies, where authoritarianism was associated with: (*a*) perception of models as strong moral models, and (*b*) strong identification with the group to which the parents belong. It appears that this researcher may have isolated a xenophilic syndrome among Jews where the parent and the reference group is Jewish. This interpretation allows us to look upon the study, therefore, as suggestive, but indirect data, along with the California studies, for the hypotheses regarding differential socialization factors producing ethnocentrism or xenophilia.

At least one precaution is needed in interpreting our hypotheses. Xenophilia or ethnocentrism must be defined with respect to sociological properties of the primary socializer. The anti-Semitic male from a Jewish family is, by our definition, a xenophile, whereas the anti-Semitic male from a non-Jewish family—let us say an Irish-Catholic family—may be, by our definition, ethnocentric. Similarly desire to travel abroad may be in some instances seen to represent identification with parents of foreign first-generation extraction and hence not reflect xenophilia, whereas desire to travel abroad may in other instances represent a lack of identification with parents of highly acculturated American backgrounds. It is always, in other words, necessary to make clear the reference group involved in the socialization process in order to differentiate clearly the xenophile from the ethnocentric orientation and, in some instances, the equalitarian from either.

## E. Summary

Questionnaire data on 132 males and 141 female college students indicated that those individuals who agree more frequently with statements indicating preferences for individuals and institutions foreign as opposed to domestic (called Xenophilia) tend to score higher on self-dislike items than those who usually disagree with these very items. Furthermore individuals scoring higher on items from the California test of Authoritarians also score higher on self-dislike than those who score low. The results seem to hold more strongly for the males than the female students. An attempt was made to account for this difference.

The results were also seen to be consistent with predictions made from a preliminary theory that self-dislike is a result of conflictful parent-child relationships which determine in various ways interest in ingroup and outgroup comparisons, and particularly xenophilic and ethnocentric attitudes.

## References

1. Adorno, T. W., Frenkl-Brunswik, E., Levinson, D. J., & Sanford, R. N. The Authoritarian Personality. New York: Harper, 1949.
2. Ausubel, D. P. Theory and Problems of Adolescent Development. Grune & Stratton, 1954.
3. Brodbeck, A. J. A test of some Adlerian hypotheses regarding superiority and inferiority in an intact group. *J. Soc. Psych.*, (in press).
4. Parsons, T. Essays on Sociological Theory. New York: Free Press, 1949.
5. Perlmutter, H. V. Relations between the self-image, the image of the foreigner, and the desire to live abroad. *J. of Psychol.*, (in press).
6. ———. Some characteristics of the xenophilic personality. *J. of Psych.* (submitted for publication).
7. Sarnoff, I. Identification with the aggressor: Some personality correlates of Anti-Semitism among Jews. *J. Pers.*, 1951, **20**, No. 2.
8. Sullivan, H. S. Interpersonal Theory of Psychiatry. New York: Norton, 1953.
9. Whiting, J. W. M., & Child, I. Child Training and Personality. New Haven: Yale Univ. Press, 1953.

*Department of Psychology*
*Boston University*
*Boston 15, Massachusetts*

*Center for International Studies*
*Massachusetts Institute of Technology*
*Cambridge 39, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 281-289.

# COMMUNICATION TO HIGH AND LOW POWER FOREIGN AND DOMESTIC AUTHORITIES*[1]

HOWARD V. PERLMUTTER AND BERNARD HYMOVITCH

*Center for International Studies, Massachusetts Institute of Technology*

## A. INTRODUCTION

What are the psychological forces that operate when an individual of one nationality is required to communicate with another nationality? What are the social psychological consequences of a foreign versus a domestic authority's potentiality to reward and punish the subject of power?

The literature in social psychology contains little but conjecture concerning the comparison of such intra- and inter-national interactions. In fact, little controlled research has been done using cross-national variables.

We have chosen, then, for this first exploratory study two variables central to the above problems and, we believe, focal in any theory of mediation of information *across* national boundaries.

The first, termed *Power* (in a social psychological sense), is defined as the capacity of one individual or group to reward and punish another individual and group (5).

The second, termed *Nationality,* is defined as a reference group property of an individual connecting him to a country,

## B. THE POWER VARIABLE

Generally, and in accordance with common observation, it is presumed that the greater the power to which an individual is subject from another, the greater his need to send communications designed to please the source of power. In some contexts, agreeing with the source of power will be seen to be pleasing, in others as displeasing.

Hymovitch and Stampolis (2) have found that when perceived power of supervisors over workers increased, the workers' readiness to communicate disagreements on the job decreased. The present study had as an objective to determine the relationship between disagreement and power in the class-

---

*Received in the Editorial Office on June 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This is one of a series of experimental researches on international communications being conducted in the International Communications Program at the Center for International Studies, Massachusetts Institute of Technology.

room situation. For, it seemed possible that the meaning of such communication might be very different in the academic atmosphere of the student-professor relationship.

## C. The Nationality Variable

Few theories are available concerning the effect of difference in nationality on the content of communications. We know at the outset from national stereotype studies that individuals have varying images of typical Americans, French, Germans, etc. (3). But how do these "mental pictures" determine in the concrete situation the quality and quantity of communications between them?

From common sense observation we might expect that: (*a*) The *prestige* of a country with respect to specific attributes may act as a "power increment" or "decrement." That is, if a person is communicating to a French art critic, he may perceive this art critic's response as more or less salient than an American critic's, e.g., in the field of art, the French critic, *in being French,* may have more capacity to reward and punish than an American art critic. (*b*) The *ingroup-outgroup effects.* Individuals experience various degrees of identification with their country. In most circumstances, communication to a foreigner may be characterized as communication from an ingroup member to an outgroup member. The ingroup member is likely to perceive the outgroup member as one whose sympathies, values, and goals are less coöperatively interdependent with his own sympathies, etc., than those of a fellow ingroup member.

One of the most important ingroup-outgroup effects concerns the kinds of information ingroup members are willing to give to outgroup members. On the international level, what attitudes and characteristics attributed to one's country does an individual report to foreign, as opposed to domestic audiences?

With these problems in mind, we undertook to pretest the differential effects of power and nationality, in a limited context, and with readily available subjects.

## D. Procedure

Forty-seven adults in a course in Human Relations at Boston University had been randomly assigned to four groups of about 11 in number (in four different rooms) and were subjected to four experimental conditions. In all of these conditions the students were asked to evaluate a series of criteria for "good groups." The students had already met weekly during a period of eight weeks for two-hour discussions on the psychology of groups. They had received no grades, and some general concern had been expressed regarding "how and when they would be graded." This, then, was the first task in which any rating of their work was made.

The criteria which the four groups were each asked to evaluate were supposed to have been developed by a professor, to whom their evaluations would be communicated. By varying the characteristics of the unnamed professor as described to each of the four groups, nationality and power were controlled in these ways: (*a*) *High power American professor:* Subjects were told that an American professor was to *grade* their evaluation of his "criteria for a good discussion group." (*b*) *Low power American professor:* Subjects were told that an American professor was to look at, but *not grade,* their evaluations of his "criteria for a good discussion group." (*c*) *High power French professor:* Subjects were told that a French professor was to *grade* their evaluations of his "criteria for a good discussion group." (*d*) *Low power French professor:* Subjects were told that a French professor would look at, but *not grade,* their evaluations of his "criteria for a good discussion group." It was assumed that the four groups were matched for attitudes toward the criteria, that the differences in their judgments were attributable to the variations in experimental conditions.

The six criteria of a good group were as follows: (*a*) A good group is one in which differences of opinion and non-conformity are encouraged and rewarded. (*b*) A good group is one in which hostility and animosity are both expected and freely expressed. (*c*) A good group is one where group action is not undertaken until every member is convinced of its wisdom. (*d*) A good group is one where each person's opinion is equally considered in making group decisions, regardless of the great differences in merit of each person's contribution. (*e*) A good group is one that would rather dissolve itself than force continual compromises on its members. (*f*) A good group is one in which leadership is more concerned with membership morale than with group productivity.

It was hoped that the six criteria would be seen as dogmatic and controversial. The subjects were to evaluate each criterion through the use of a seven-point scale. This seven-point scale of agreement allowed for a range of opinions from 1: disagree completely, to 7: agree completely.

The students set to work vigorously on the two phases of the experiment: (*a*) First each student individually rated each of the six criteria along the scale of agreement and added qualifying comments (messages to the source of power). (*b*) Then all students discussed each criterior with the other members of the group in order to reach *one collective* rating for each group on each criterion. (*c*) Finally, all students as individuals again filled out an attitude questionnaire which was introduced in this way: "Professor ——— would like your opinions on the following issues in order to assess some of the common interests of the members of your group."

Section (*a*) of the questionnaire required the students to judge the likelihood that each of a series of statements had been made in a discussion group by "either an American or a Frenchman." The statements were borrowed from the Frenkl-Brunswik test of Authoritarianism, e.g., "No obstacle can prevent a person from reaching his goal, providing he is determined enough."

Then, in Section (*b*), students rated their own opinion of the authoritarian statements on a seven-point scale of agreement.

## E. RESULTS AND DISCUSSION

### 1. *Communicated Evaluation of Source of Power's Opinions*

The evaluations (on the seven-point scale) by all members of a group of all "The Six Criteria for a Good Group" were averaged. Table 1A summarizes the results.

TABLE 1A

MEAN RATINGS OF THE "SIX CRITERIA OF A GOOD GROUP"* COMMUNICATED TO POWER FIGURE BY FOUR EXPERIMENTAL GROUPS

| Power figure presented as | | | |
|---|---|---|---|
| Low power American | High power American | Low power French | High power French |
| 4.5 | 4.3 | 4.8 | 4.2 |

*The higher the mean, the greater the agreement with the authority figure.

Although the differences in the means between high and low power groups are not statistically significant, the data suggest that the individuals in groups over whom the authority had little power tended to agree with his product to a slightly greater extent than the groups over whom the authority had considerable power.

A conservative interpretation of the data would at least state that we have found no evidence for an increase in agreement with increase in power as has been found in other research.

The means of the collective evaluations reached by each of the four groups in a discussion of each of the six criteria is contained in Table 1B.

Although the data are meagre for statistical analysis, the results again support the finding that more agreement was communicated to low power figures than to high power figures.

TABLE 1B
MEAN "GROUP QUA GROUP" RATINGS OF "SIX CRITERIA OF A GOOD GROUP"*

| Power figure presented as | | | |
|---|---|---|---|
| Low power American | High power American | Low power French | High power French |
| 4.33 | 3.50 | 5.33 | 3.00 |

*The higher the mean, the greater the agreement with the authority figure.

## 2. *Qualitative Agreement*

A further test of the amount of agreement with the authority figure was available through an analysis of the messages that accompanied the quantitative evaluations. The messages sent by each person were coded along a seven-point scale of agreement-disagreement. An average score along the dimension was then obtained for *each* individual. Table 2 presents the results of that analysis. Intercoder agreement on ratings was approximately 85 per cent.

TABLE 2
NUMBER OF PERSONS WHOSE COMMUNICATIONS WERE IN AGREEMENT OR DISAGREEMENT WITH AUTHORITY'S "CRITERIA OF A GOOD GROUP"

| Groups with | Persons in agreement | Persons in disagreement |
|---|---|---|
| Low power authority | 14 | 3 |
| High power authority | 6 | 13 |

$X^2$ (Yates Correction) = 4.71.
$P = < .03$.

These data, then, strengthen the suggestion implicit in the communicated ratings (Tables 1A and 1B) that the groups under a high power figure tend to agree less with the authority figure than those under a low power figure.

It would seem, then, that under certain conditions "the desire to please the power figure" may take the form of disagreement, rather than agreement. Indeed, the academic atmosphere of our experiment may very likely suggest that greater rewards come from critical evaluation than from blind acceptance.

It is, of course, possible that disagreement may still represent displeasure in the classroom situation. If true, then one could interpret these data as suggesting less desire to please the high power figure than the low power figure. This interpretation would then be contrary to the findings of many studies (1, 3, 4, 6).

## 3. *Quantity of Communications*

Each individual had the opportunity to send a maximum of six written communications to the authority figure. However, they were not compelled to write six communications. A "blank" message may be interpreted as due to low motivation to communicate or to a caution of communicating. Considering the facilitation and encouragement to communicate in the experimental format, low motivation to communicate may be interpreted as a result of fear that one will say the wrong thing—a likely response to an ambiguous, threatening situation. For this reason, we would expect that more blank messages should be sent in the high power situation. The results seem consistent with this interpretation.

TABLE 3

NUMBER OF PERSONS WHO COMMUNICATED TWO OR MORE VERSUS NONE OR ONE BLANK MESSAGE WITH RESPECT TO AUTHORITY'S "CRITERIA OF A GOOD GROUP"*

| Groups with | Persons with two or more blank messages | Persons with none or one blank message |
|---|---|---|
| Low power authority | 5 | 16 |
| High power authority | 15 | 7 |

$X^2$ (Yates Correction) $= 4.14$.
$P = < .05$.

*This division was made by dichotomizing the scores of the total population including both experimental groups.

### 4. *Statements About Americans and Frenchmen*

A series of statements was presented to the subjects who were asked to decide whether each statement was most likely to have been made by an American, Frenchman, or both. The subject was later asked whether he himself was favorably disposed toward the behavior described in each statement. As a result, it was possible to determine: (*a*) The number of statements attributed to Americans with which the subject agreed; (*b*) the number of statements attributed to Frenchmen with which the subject agreed; (*c*) the number of statements attributed to Americans with which the subject disagreed; (*d*) the number of statements attributed to Frenchmen with which the subject disagreed.

By subtracting the score each subject made of (*c*) from the score made on (*a*), we could determine the relative agreement with statements attributed to Americans ($a — c$).

In the same way we could obtain the relative agreement with statements attributed to Frenchmen ($b — d$).

Further, by subtracting the number of resulting scores with Frenchmen ($b — d$) from the resulting score with Americans ($a — c$), we could obtain a score that would describe the differential agreement by nationality.

A positive score would here indicate more agreement with statements attributed to Americans than with statements attributed to Frenchmen.

Thus, comparisons were possible on the above-described scores in the four experimental conditions.

There were no significant differences between groups in high and low power situations on any of these scores. The results of a comparison between the two groups who were communicating to an American versus the two groups who were communicating their opinions to a Frenchman were clearer.

The mean score of Agreement-Disagreement attributions to Americans ($a - c$) for the two groups communicating to an American was —.8 (the minus sign indicates more disagreement with attributions to Americans than agreement). For the two groups communicating to a Frenchman the mean score was +.7, indicating more agreements with statements attributed to Americans than disagreements. The *T*-test was run and *P* was found to be a .06 level of significance. Thus, this finding suggests that these American subjects were more likely to describe Americans in a favorable light when

These results are consistent with those obtained by subtracting attribution to Frenchmen from attribution to Americans.

The average scores of "Agreement-Disagreement" attributions to Frenchmen for the two groups communicating to an American was +.95 while for the groups communicating to a Frenchman the mean was +2.80. The *P* value from a *T*-test was significant at less than the .07 level of significance. It should be noted that a comparison of attributions to Americans versus attributions to Frenchmen by each group yielded consistent results favoring the Americans—as manifest by the fact that both of these latter means are positive.

### 5. *The Nationality Variable*

Our results then indicate some evidence for the ingroup-outgroup effect of the nationality variable. If the ingroup member perceives the outgroup member (a foreigner) as less inter-dependent with his own sympathies, etc., than a fellow ingroup member, it is likely that he will try to provide a better picture of his ingroup to the outgroup member than to another ingroup member. The nationality variable, it appears, can induce psychological situations not too different from those in other ingroup-outgroup relations. When one's country becomes the subject's ingroup, and another country the outgroup, these factors evidently arise.

This does not explain why these ingroup-outgroup effects did not "take" in the first phase of the experiment. It seemed less important that the source of power was either French or American, but more important that he could reward or punish. Several students remarked when the experiment was explained later that they felt that there was no reason to believe that a French professor of psychology would believe very differently from, or be more or less capable than, the American professor. This would suggest that under some conditions the nationality variable is less important than other variables. Indeed, later research has indicated that the college students' image of a typical French professor correlates significantly highly with the image of a typical American college professor.[2]

## F. SUMMARY

The subjects were 47 adults in a Human Relations course who had been randomly assigned to four groups. They had met weekly for eight weeks prior to the experiment. Students in the different groups were required to evaluate individually and as a group the "Criteria for a Good Group" that a professor had developed. In the high power situation they were told that an American (or French) professor was going to grade their evaluations. In the low power situation they were told that an American (or French) professor was going to look at but not grade their evaluations.

The students in all four groups were asked to communicate their attitudes toward American and France to the power figure. Results were as follows:

1. The greater the source of power (whether foreign or domestic), the greater was the disagreement with the source of power in the academic atmosphere.
2. Significantly more blank messages were sent to high sources of power.
3. Americans were consistently described in a more favorable manner than Frenchmen, regardless of the nationality of the power figure.
4. Americans were described more favorably when the students communicated to the French power figure than when they communicated to the American power figure.

---

[2] Research completed with Professor J. Bruner at Harvard University.

## REFERENCES

1. HURWITZ, J. I., ZANDER, A. F., & HYMOVITCH, B. Some effects of power on the relations among group members. In *Group Dynamics* (*Ed.* Cartwright, D., & Zander, A. F.). New York: Peterson, 1953.
2. HYMOVITCH, B., STAMPOLIS, A., & BERKOWITZ, N. Communication in an industrial hierarchy. (In preparation.)
3. KATZ, D., & BRALEY, K. Racial stereotypes of one hundred college students. *J. Abn. & Soc. Psych.*, 1933, **28**, 280-290.
4. LIPPITT, R., POLANSKY, N., REDL, F., & ROSEN, S. The dynamics of power. *Human Rel.*, 1952, **5**, 37-64.

5. PERLMUTTER, H. V. Impressions of influential members of discussion groups. *J. of Psych.*, 1954 (in press).
6. ZANDER, A. F., COHEN, A. R., HYMOVITCH, B., & STOTLAND, E. Some Determinants of Rôle Relations. Ann Arbor, Mich.: Institute for Social Research, 1952 (mimeo).

*Center for International Studies*
*Massachusetts Institute of Technology*
*Cambridge 39, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 291-300.

# SOME CHARACTERISTICS OF THE XENOPHILIC PERSONALITY*[1]

*Center for International Studies, Massachusetts Institute of Technology*

HOWARD V. PERLMUTTER

## A. INTRODUCTION

When Henry James expressed his sentiments through the artist-hero of *The Madonna of the Future,* he represented a point of view of the alienated American intellectual of the 1920's. He wrote:

> We are the disinherited of art. We are condemned to be superficial. We are excluded from the magic circle. The soil of American perception is a poor, little, barren and artificial deposit. Yes! We are wedded to imperfection. An American to excel has just ten times as much to learn as a European. We lack the deeper sense: we have neither taste, nor tact, nor force. How should we have them? Our crude and garish climate, our silent past, our deafening present, the constant pressure about us of unlovely circumstances are as void of all that nourishes and prompts and inspires the artist as my sad heart is void of bitterness in saying so. . . .

Strangely, the study of such alienated individuals, or more particularly, the study of champions of things and people "foreign," has received little attention. The intellectual energies of our time seem leveled primarily on the "ethnocentric typology," in contrast with a favored "equalitarian syndrome" (1, 3). This applies generally to anthropological, sociological, and psychological studies.

What in previous research on personality and social psychology might provide us with potential evidence on the Xenophile and his relation to the other types studied?

The evidence is still growing that the attitude-syndrome of the low Authoritarians is by no means homogeneous. The investigators were in fact explicit that their focus was the anti-democratic personality as opposed to all others. This is clear after an examination of the item content. The high agreers or high authoritarians of the *F*-scale were expressing approval of a

---

*Received in the Editorial Office on June 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This is one of a series of studies of "Personality Determinants of International Communications" being conducted in the International Communications Program at the Center for International Studies, Massachusetts Institute of Technology.

series of relatively dogmatic statements, which, content-wise, illustrate the authoritarian syndrome. A list of these dimensions would include:

1. A very conventional outlook on "life."
2. An identification with the strong.
3. A denial of impulses.
4. An attribution (projectivity) of badness to foreigners.
5. A hatred of foreigners.
6. A submission to government.
7. A belief in organization.
8. Conservatism in politics.
9. Interest in material security.
10. Anti-intraception (psychological approach to others).
11. Inability to rebel against parental figures.

It was more or less assumed that individuals who score *low* on these dimensions are necessarily equalitarians, but it is clear that if an individual disagrees with the item:

> "The American way of life is the best way of life,"

he may be thinking of at least *two very different alternatives.*

1. America has a way of life suitable for Americans: Europeans have have their own suitable point of view.

*Or*

2. America's way of life is assuredly *inferior* to the foreign way of life.

In both cases a person could score low, yet in one case, 1, a flexible, tolerant outlook is expressed, while in the other, 2, dichotomization and stereotyping of properties of in- *versus* out-group members.

Logically, at least, it is possible to draw a list of characteristics of individuals which represent the converse of the anti-democratic criteria, but do not *represent* what has come to be known as an equalitarian outlook. Schematically such a Xenophilic-anti-authoritarian type would display strong dispositions toward:

1. Being very unconventional.
2. An identification with the weak.
3. A glorification of his impulses.
4. Attribution (projectivity) of badness to Americans, goodness to foreigners.
5. Love and admiration for foreigners, generalized widely.
6. Hatred for government or authority.
7. Belief in disorganization, anarchy.
8. Radicalism in politics.

9. Strong antipathy against an interest in material security.
10. A deeply psychological approach to others and oneself.
11. An emphasis on rebellion toward parental figures.

For a first approach we shall be primarily concerned with the Xenophile high on Items 4 and 5. Whether the syndrome described above exists as a total psychological reality remains a moot question. Actually it is our thesis that Xenophiles who strongly exhibit Characteristics 4 and 5 would possess *underlying* attributes similar to the Authoritarian, depending on the subtlety of the questions posed to him. So a Xenophile can have at the same time a preference for foreign objects, etc., and score high on suspicion of others, particularly authority figures, belief in the intrinsic evil of human nature, the dichotomization of the rational and the emotional, and the use of projection in personal relationships.

## B. KINDS OF XENOPHILES

Xenophilia means literally a love for strangers and foreigners. We added to this definition a further criterion: ". . . an implicit or explicit disrespect for or hatred of one's own sociological reference group."

There are many kinds of Xenophile, born perhaps of a variety of social contexts and with different situational and genetic determinants.

The Xenophile may be:

1. A person who dreams of living abroad in the exotic lands of great pleasure.
2. An intellectual, acculturated in European tradition and values and identified with European life.
3. An advocate of world government who believes fervently in the disasters of nationalism and the necessity for identifying with a larger super-political entity.
4. A Communist who is allied to some cause led and directed by a foreign power.
5. An artist, rebellious and unconventional in situational or permanent adolescence.
6. A second generation individual, unacculturated and unassimilated, who still maintains primary ties with the old world.

The assumption is a core orientation toward the in-group and out-group operates in the Xenophile and underlies the variety of apparently heterogeneous motives for consistently liking or admiring things and people foreign.

## C. MAJORITY-MINORITY RELATIONS: THE AUTHORITARIAN AND THE XENOPHILE

Let us examine in some detail a few of the characteristics of a Xenophile as derived from the list discussed above.

One such property of the Xenophile is his apparently strong identification with a minority individual or group. According to our conception, he says he feels deeply connected with the suffering and the subordinate, the underprivileged, the objects of prejudice or economic inequities.

The authors of *The Authoritarian Personality* suggest that:

> 1. An individual may be concerned with many groups which are neither in-groups or out-groups for him.
>
> 2. An individual may feel sympathetic toward Negroes *without identifying with them,* or conversely, may be opposed to many groups in the sense of feeling a difference in interests and values, or merely feeling that their aims and existence are irrelevant to him . . . but these are not out-groups if there is not the sense of contra-identification.

This phenomenon of sympathy without identification is supposedly characteristic of the equalitarian personality. The high authoritarian more frequently thinks in terms of in-groups and out-groups. If he cannot identify, he must oppose; if a group is not acceptable, it is alien.

*We would expect a similar kind of dichotomous thinking in the Xenophile.* The Xenophile may identify with the minority to work out his aggressions toward the majority or "the authority-group." He, like the Authoritarian, is concerned with the authority. His feelings toward it become expressed in a somewhat indirect way. The intellectual Xenophile has no doubt a layer of qualificatory expressions, socially acceptable, but we should expect areas of rigidity similar to those in the Authoritarian.

## D. IDEAS CONCERNING HUMAN NATURE: THE XENOPHILE AND THE AUTHORITARIAN

The original study of the Authoritarian personality further pointed out that the Authoritarians find it difficult to identify with *humanity,* and that their resolution takes the form of nationalism and cynicism about world government and peace. It takes other forms, too, such as:

> 1. Ideas concerning the intrinsic evil of human nature.
>
> 2. Ideas that evil is unchangeable, rationalized by pseudoscientific hereditarian theories of nature.
>
> 3. Ideas that human nature is intrinsically aggressive, lazy and power seeking.

The authors of *The Authoritarian Personality* insist that the democratic alternative "is not a vague and abstract 'love for everybody' but ability to like and to dislike, to value and oppose individuals on the basis of concrete and specific experience." This may be the democratic alternative, but it is not the only alternative. We would suspect that the Xenophile may often be the lover of "abstract humanity" or the conscious humanitarian and unconscious hater of the particular individual.

The Xenophile may in fact feel that there are all kinds of inexplicable "goings on," and that man has as an intrinsic part of him evil, aggressive impulses. However, he would work out his concern with his impulsive life through his own basic defensive process: Projecting the good, the desirable, the beautiful onto things foreign—a relatively persisting "grass-grows-greener-on-the-other-side orientation. *This again does not differ too greatly from the Authoritarian.*

## E. The Problem

Our general propositions, to sum up the previous, would be:

*Xenophilia is based on an unconscious but rigid in-group-out-group distinction; it involves stereotyped, negative imagery and hostile attitudes regarding in-groups, stereotyped, positive imagery regarding out-groups (often where he has had no direct experience with that out-group). The underlying dynamics of the Authoritarian and Xenophile are probably very similar.*

This present study was aimed at comparing for a limited sample of individuals some of the properties of the Xenophile and the Authoritarian. Our prediction, contrary to common sense, would be: Individuals scoring higher on Xenophilia should score significantly higher on Authoritarianism than individuals scoring low on Xenophilia.

## F. Devising of the Xenophile Scale

Since no test for Xenophilia was available, it was necessary to develop and pretest a set of items coördinated with the dimensions of Xenophilia mentioned above.[2] The Frenkel-Brunswik procedure was followed and a series of items high in discriminating power (evoking a range of opinions) was found. The items discussed in this paper was aimed at tapping preferences of foreign objects, versions, and institutions over American. Usually, they involve relatively simple and dogmatic comparisons. For this reason, it is unlikely that the more sophisticated Xenophile, who probably does not categorically agree with anything stated there, was reached. But since the population tested were to be college sophomores and juniors, the items had to be clear and unequivocal. As the numbers indicate below, the Xenophile items were interspersed with other questions so that the purpose of the test was not apparent.

---

[2] Test-retest reliability over a period of one month for a mixed sample of 30 subjects was .85.

The instructions were:

> We are trying to find out what the general public feels and thinks about a number of important questions.
>
> This is *not* an *intelligence test* nor an information test. *There are no "right" or "wrong" answers.* The best answer is *your personal opinion.* You can be sure that, whatever your opinion may be in a certain question, there will be many people who agree, many who disagree. Similarly, you will probably find yourself agreeing strongly with some statements, disagreeing just as strongly with others, and being perhaps more neutral about still others. Please:
>
> (*a*) Read each statement carefully and mark it according to your first reaction. It isn't necessary to take a lot of time for any one question.
>
> (*b*) *Answer every question.*
>
> (*c*) *Give your personal point of view.*
>
> (*d*) Be as *sincere, accurate* and complete as possible in the limited time and space.
>
> Please mark each statement in the left margin, according to the amount of your disagreement or agreement, by using the following scale:
>
> +1: slight support, agreement
> +2: moderate support, agreement
> +3: strong support, agreement
>
> ---
>
> —1: slight opposition, disagreement
> —2: moderate opposition, disagreement
> —3: strong opposition, disagreement

*Xenophile Sub-Scale Items*

3. The chief stimulants to basic American institutions in this country have come mainly from European and Oriental ideas and doctrines.
4. The most delicious foods are European or Oriental.
11. The British use the English language better than most Americans.
13. European men are usually more romantic than American men.
15. Too much stress is placed on Americanism these days, not enough on the fact that we are all descendants of the Old World.
17. Compared to the French, Americans are an unimaginative people.
24. European children are generally better mannered than Americans.
25. Europeans on the whole appreciate and understand the arts better than Americans.
26. Most European girls make better wives than American girls.

27. Europeans generally are a warmer and friendlier people than Americans.

Those who scored high on these scales, indicating relatively high agreement with these items, were called High Xenophiles (hi-XPLS).

In addition, all students filled out this series of items from the Levinson-Frenkl-Brunswik test:

### *Authoritarian Scale*

28. America may not be perfect, but the American way has brought us about as close as human beings can get to a perfect society.

29. Men like Henry Ford and J. P. Morgan, who overcame all competition on the road to success, are models for all young people to admire and imitate.

30. In general, full economic security is bad; most men wouldn't work if they didn't need the money for eating and living.

31. No sane, normal, decent person could ever think of hurting a close friend or relative.

32. Human nature being what it is, there will always be war and conflict.

33. No weakness or difficulty can hold us back if we have enough will power.

34. People can be divided into two distinct classes, the weak and the strong.

35. It's a mistake to trust anyone who doesn't look you straight in the eye.

36. Most people don't realize how much our lives are controlled by plots hatched in secret by politicians.

37. Some day it will probably be shown that astrology can explain a lot of things.

38. Familiarity breeds contempt.

39. The true American way of life is disappearing so fast that force may be necessary to preserve it.

40. The wild sex life of the old Greeks and Romans was tame compared to some of the goings on in this country, even in places where you might least expect it.

41. Sex crimes, such as rape and attacks on children, deserve more than mere imprisonment; such criminals ought to be publicly whipped or worse.

42. Every person should have complete faith in some supernatural power whose decisions he obeys without question.

43. Homosexuals are nothing but degenerates and should be severely punished.

44. Nowadays when so many different kinds of people move around and mix together so fully, a person has to be especially careful to protect himself against infection and disease.

45. It is possible that wars and social troubles will be ended once and for all by an earthquake or flood that will destroy the whole world.

## G. RESULTS AND DISCUSSION

According to our theory, individuals scoring high on a test of Xenophilia would score relatively higher on Authoritarianism than individuals scoring low on Xenophilia. Here the Xenophilic items emphasized preference for foreign objects, persons, institutions over their domestic counterparts. Our Authoritarian items weighted heavily such tendencies as suspicion of others, dichotomous, stereotypic thinking, intrinsic evil of human nature, belief in mystical forces.

The items described above were administered to two groups of students at Boston University. One sample consisted of 140 females at the School of Practical Arts and Letters; the other of 130 males at the College of Business Administration.

The scores of all subjects were translated to a scale of seven so that strong support or agreement (rated +3 by the subject) was given —7 by the scorer, while strong opposition or disagreement (rated —3 by the subject) was given a —1 score. Our results will be confined to the 10 Xenophile items, and 17 Authoritarian items described above. The subjects also filled out other scales which will be discussed in future publications.

A group of approximately the top and bottom 20 of the distribution were selected from the male and female groups.[3]

The top group was called High Xenophiles, and the bottom group was called Low Xenophiles. Then, the corresponding Authoritarian scores were determined for each of these individuals, and a *T*-test for significance of the difference between the means of the two distributions was calculated.

Table 1 shows that the hypothesis stated above is supported by the data. The High Xenophiles do in fact score significantly higher in Authoritarianism than the Low Xenophiles. This difference is significant at the .01 level of significance for both males and females.

The mean *A*-score of 4.0 for males and females is slightly in the Authoritarian direction, since it is greater than the mean *A*-scores of the respective distributions, and the mean *A*-score of the group scoring in the middle range of the *XPL* distribution. But since the mean *A*-score is not greater than 4.0, it does not follow that High *XPL*s are also High Authoritarians. For the same reason, the evidence does not show that the High Xenophiles do not show a rigid tendency to respond positively to items on the *A*-test. They do, of course, respond more positively to *A*-test items than Low Xen-

---

[3] Since there were considerably more scores on the low Xenophile side, care had to be taken to choose the highest Xenophile scores, whose mean item score was greater than 4.0.

TABLE 1

COMPARISON OF AUTHORITARIAN SCORES FOR HIGH, MIDDLE, AND LOW XENOPHILES

| | *N* | Mean *XPL* per item | Mean *A*-score per item |
|---|---|---|---|
| | *CBA males, N* = 132 | | |
| Hi *XPL* | 29 | 4.9 | 4.0 |
| Mid *XPL* | 69 | 3.4 | 3.5 |
| Lo *XPL* | 34 | 2.4 | 3.2 |
| Over-all averages | 132 | 3.5 | 3.6 |
| | *PAL females, N* = 141 | | |
| Hi *XPL* | 30 | 5.1 | 4.0 |
| Mid *XPL* | 78 | 3.7 | 3.7 |
| Lo *XPL* | 33 | 2.5 | 3.3 |
| Over-all average | 141 | 3.7 | 3.7 |

ophiles.[4] In this respect, we may be getting an effect similar in nature if not in degree to Cohn (2). We would prefer to conclude: the data suggest that those individuals who consistently prefer foreign objects, etc., to domestic objects tend to be more oriented toward the aspect of Authoritarian syndrome that stresses properties such as dichotomous, stereotypic thinking, suspicion of others, and belief in the intrinsic evil in human nature.

## H. THE EQUALITARIAN AND THE XENOPHILE

Since the statistical differences refer to the means of the respective distribution, it is clear that some Xenophiles score relativly low in Authoritarianism and some, necessarily the average, score very high. The results suggest, then, different "breeds" of individuals cross cut the current notion of the rigid, conservative, ethnocentric, high authoritarians, and the flexible, democratic, internationalistic, low authoritarians.

It is obvious, moreover, that further clinical and experimental research is needed to compare the motivational-cognitive-affective underpinnings of these typologies.

## I. CONCLUSION

We started out postulating what seemed to be a logical converse of the Ethnocentric-Authoritarian syndrome, namely, a Xenophilic-Anti-Authoritarian syndrome. In focusing on two aspects of this constructive syndrome, particularly those individuals who project badness on Americans, goodness on foreigners, and who generalize widely their admiration for foreign institutions, people, and objects, we were led to a theory that predicted some basic underlying similarities between the Xenophile and Authoritarian personality. The results of Xenophile and Authoritarian questionnaires administered to 133 males and 147 female students at Boston University produced the following result:

---

[4] Future researches will control this positive response set by varying the manner of presentation of the items and including converse content items.

1. Fifty-nine individuals scoring highest on the Xenophile scale scored, on the average, significantly higher on the Authoritarian scale than 67 individuals scoring lowest on the Xenophile scale.
2. The mean Authoritarian score of the High Xenophiles was higher than the mean Authoritarian score of the entire distribution of subjects.

## REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. The Authoritarian Personality. New York: Harper, 1949.
2. COHN, T. S. The relation of the F-scale to a response to answer positively. *Amer. Psychol.*, 1953, **8**, 335 (abstract).
3. ROKEACH, M. Generalized mental rigidity as a factor in ethnocentrism. *J. Abn. & Soc. Psychol.*, 1948, **43**, 259-278.

*Center for International Studies*
*Massachusetts Institute of Technology*
*Cambridge 39, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 301-312.

# THE NON-VOTER IN 1952: A STUDY OF PITTSFIELD, MASSACHUSETTS*

*Department of Psychology, Williams College*

Philip K. Hastings

The fact that a large segment of the American electorate fails to exercise its voting privilege and responsibility on election day is of both theoretical and practical concern. In the presidential elections since 1924, between 35 and 45 per cent of the eligible voters have failed to go to the polls; the percentages are even higher in the non-presidential or off-year elections. Although recognition of this political fact has resulted in a variety of tentative speculations, there is a paucity of empirical research pertinent to the problem. The purpose of this paper is to present research findings, based on a 1952 community study (2), relevant to the thesis that level or degree of voting participation constitutes but one manifest indication of the degree to which an individual is both psychologically and behaviorally involved in the political processes of his community.

A significant proportion of political behavior research dealing with non-voting has been limited essentially to a sociological description of the non-voter group. It has been found repeatedly, for example, that voting participation decreases as one moves from high to low on economic, educational, and occupational continua. Further, a higher frequency of non-voting is apparently characteristic of women, Negroes, and Democrats (4-6). Other studies have attempted to correlate level of voting participation with the distinctive characteristics of given election situations. For example, is the contest generally considered to be a runaway, or is there considerable doubt regarding the outcome? To what degree have the party organizations extended themselves? Have the dominant political issues caught the fancy of the voting public? Important as these data are, they are primarily descriptive in nature, and do not provide much insight into the reasons for failure to vote.

While each of the various relationships established in the research cited above stands in its own right as a measurable contribution to our understanding of the non-voting problem, it is the writer's opinion that in the aggregate, these findings assume added significance if viewed in light of the thesis that

---

*Received in the Editorial Office on June 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

the non-voter group constitutes ". . . a more or less socially isolated segment of the population. Non-voting is merely one manifestation of the weak linkage of a category of the population to social affairs generally" (4, p. 580). It is within the context of this theme that the empirical data presented in the following pages should be viewed.

In analyzing the Pittsfield non-voter group, two primary questions were posed. First, is failure to vote an isolated behavioral phenomenon, or is it only one example of a more general tendency to avoid overt participation in various facets of the social life of one's community? The empirical data presented in the following pages support the latter alternative. The results of the 1952 election study of Pittsfield, Massachusetts, indicated a positive correlation between amount of voting participation and amount of (*a*) voluntary group membership, (*b*) exposure to political communication, and (*c*) acquired political information.

Our second question asked: are there psychological correlates to the behavioral fact of chronic non-voting? Positive correlations were established between amount of voting participation and (*a*) frequency of self-identification with political groups, i.e., the three main political divisions: Democrat, Republican, and Independent, (*b*) frequency with which assumptions were made regarding the direction of voting behavior of individuals composing various politically stereotyped social groups, and (*c*) a negative relationship was established between level of voting participation and the frequency with which the Pittsfield respondents expressed no opinion regarding a number of 1952 political campaign issues.

In each of the four sample surveys[1] conducted in Pittsfield during the period April through November, the respondents were requested to indicate both if and how they had voted in various 1948, 1950, and 1952 (November survey) political contests. Using these data, we grouped each individual according to his level of voting participation. Three respondent categories were established. The chronic non-voting group included those individuals who said that they had not voted in any of the 1948, 1950, and 1952 contests about which they were questioned.[2] Their counterparts, classified as consistent voters, claimed to have voted in all of the political races. A third group contained the remaining respondents, all of whom indicated that they had exercised their voting privilege in at least one of the various contests about which they were queried.

---

[1] Probability samples were drawn. The sample sizes and confidence limits are included in footnote 2 of (3).

[2] (1) 1948: Presidential.
(2) 1950: Gubernatorial, congressional.
(3) 1952: Presidential, senatorial, gubernatorial, congressional.

As has been indicated above, one facet of the central theme of our analysis of the Pittsfield non-voter was that failure to go to the polls on election day is not an isolated phenomenon, but rather is consistent with a general pattern of low participation in community affairs. What was the nature of the evidence relevant to this question? Among the specific variables considered to be relevant to this assumption were group membership behavior, exposure to political communication, and acquired level of information. Our concern was not with the direction or specific content of the individual's behavior in these areas, but instead, focused upon the degree or amount of participation characteristic of him.

Specifically, our initial hypothesis held that there exists a positive relationship between amount of voting and amount of voluntary group membership behavior. In April, August, and October the survey respondents were asked to give the names of three organizations or groups (other than their family) of which they were members.[3] As will be noted in Table 1A, the obtained data supported the hypothesis. As level of voting participation increases, so does the percentage of individuals claiming at least one voluntary group membership. The percentage difference between the chronic non-voter and

TABLE 1
VOLUNTARY GROUP MEMBERSHIP BEHAVIOR

| | Non-voter | Occasional voter | Voter |
|---|---|---|---|
| A. Extent of Voluntary Group Membership | (*N*=111)<br>% | (*N*=115)<br>% | (*N*=524)*<br>% |
| No voluntary group membership | 52 | 36 | 29 |
| At least one voluntary group membership | (*1*) 48 | (*2*) 64 | (*3*) 71 ** |
| B. Pattern of Voluntary Group Membership | (*N*=56)<br>% | (*N*=74)<br>% | (*N*=384)<br>% |
| Church | 43 | 46 | 45 |
| Labor | 18 | 15 | 15 |
| Recreation | 36 | 32 | 41 |
| Civic and charity | 7 | 8 | 16 |
| Other | 20 | 20 | 28 |

*Throughout the four Pittsfield surveys, the non-voter groups were essentially equivalent both in terms of proportion of the total sample, and in sociological make-up. Thus, when possible, the statistics included in this article are based upon the total number of non-voter cases in the various samples. The same procedure was followed with regard to the occasional voter, and voter groups.

**Differences significant within the 5 per cent level of confidence or better include: 1 & 2; 1 & 3.

---

[3] "Would you give me the names of three organizations or groups (other than your family) of which you are a member?"

the voter groups is significant within the 5 per cent level of confidence. Analysis of the pattern of voluntary group memberships characteristic of the three respondent groups, however, revealed no significant differences (see Table 1B).

The respondent's level of exposure to political communication during the campaign months of the 1952 election was also examined as a dimension of his overt social behavior. It was hypothesized that the chronic non-voters, so designated in light of their failure to vote in the 1948 and 1950 elections, would expose themselves to 1952 political communications significantly less than the voting group. In short, it was expected that minimum voting participation and low communication exposure would be positively correlated. This relationship was tested throughout the Pittsfield community study, and as indicated in Table 2, the data consistently supported the hypothesis. It will be noted that the problem was attacked using two different approaches or types of questions. Not only were the respondents queried regarding degree of actual exposure, but also, they were asked to indicate their awareness of certain political issues that were currently being given considerable publicity. It was expected that if our hypothesis was valid, the results obtained from direct questions regarding amount of exposure, and questions dealing with awareness of central campaign issues would be consistent. This expectation proved to be well-founded. The chronic non-voters not only admitted to significantly less exposure to political communications, but further, as a group they were less aware of such major issues as the Nixon fund and corruption in government.[4] It should also be noted in Table 2 that the level of exposure characteristic of the middle group of respondents (individuals who claimed to have voted in some, but not all of the contests about which they were questioned) consistently falls between the levels of the two extreme groups. In the writer's opinion, this result further strengthens the obtained relationship.

Extent of voluntary group membership behavior, and amount of exposure to political communication may be considered relatively direct measures of an individual's contact with the social life of his community. The third overt behavioral variable (acquired level of political information) investigated in the 1952 Pittsfield research was somewhat less direct. The decision to measure level of political information, and further, to interpret the results as one index of individual-community contact, involved making the assumption that the amount of political information an individual has acquired reflects a pat-

---

[4] The items in Table 2 constitute a typical sample drawn from the four survey ballots used in the Pittsfield Project.

TABLE 2
EXPOSURE TO POLITICAL COMMUNICATION

| | Non-voter | Occasional voter | Voter | |
|---|---|---|---|---|
| A. April Survey: "Have you heard or read anything about corruption in our national government?" | | | | |
| | (*N*=49)<br>% | (*N*=46)<br>% | (*N*=247)<br>% | |
| Exposed | (*1*) 65 | 65 | (*2*) 87 | * |
| Not exposed | 35 | 35 | 13 | |
| B. August Survey: "Did you happen to listen to (or watch on television) the political conventions that were held in Chicago last month?" | | | | |
| | (*N*=32)<br>% | (*N*=32)<br>% | (*N*=145)<br>% | |
| Exposed | (*3*) 63 | 81 | (*4*) 88 | * |
| Not exposed | 37 | 19 | 12 | |
| "If exposed—would you say that you listened to (and/or watched) about ¼, ½, ¾ or all of the conventions?" | | | | |
| | (*N*=20)<br>% | (*N*=26)<br>% | (*N*=127)<br>% | |
| 1-25% | (*5*) 80 | 54 | (*6*) 42 | * |
| 26-50% | 20 | 27 | 36 | |
| 51-100% | (*7*) — | 19 | (*8*) 23 | * |
| C. October Survey: "Recently there has been some discussion about the Nixon and Stevenson special funds. Have you heard about the Nixon fund?" | | | | |
| | (*N*=30)<br>% | (*N*=37)<br>% | (*N*=132)<br>% | |
| Exposed | (*9*) 63 | 78 | (*10*) 92 | * |
| Not exposed | 37 | 22 | 8 | |
| "Now regarding the Stevenson fund,—have you heard about it?" | | | | |
| Exposed | (*11*) 47 | 54 | (*12*) 75 | * |
| Not exposed | 53 | 46 | 25 | |
| D. November Survey: "While the recent political campaign was going on, did you listen to any political speeches on the radio (and/or television)?" | | | | |
| | (*N*=15)<br>% | (*N*=28)<br>% | (*N*=141)<br>% | |
| Exposed | (*13*) 67 | 82 | (*14*) 91 | * |
| Not exposed | 33 | 18 | 9 | |

*Differences significant within the 5 per cent level of confidence or better include: 1 & 2; 3 & 4; 5 & 6; 7 & 8; 9 & 10; 11 & 12; 13 & 14.

tern of past overt behaviors characteristic of him. In light of this assumption, it seemed probable that a positive relationship would exist between low level of information and low incidence of voting behavior.

In the April and October surveys, a series of factual questions was asked. Most of these items called for an identification of political leaders ranging from the local to the national level. In addition, the respondents were asked the number of United States Senators from each state, the number of delegate votes necessary to nominate the Republican presidential candidate, and the population of Pittsfield. The items were evaluated as equal units, each with a score of one. In the April poll the maximum score was 13, while in October the highest possible tally was 10. The results presented in Table 3 show a significant relationship between the two variables: level of political information and level of voting participation.

TABLE 3

LEVEL OF POLITICAL INFORMATION

| | Non-voter | Occasional voter | Voter |
|---|---|---|---|
| A. April Survey: | | | |
| Score: | (*N*=49)<br>% | (*N*=46)<br>% | (*N*=247)<br>% |
| 0-6 (low) | (*1*) 41 | 35 | (*2*) 13 * |
| 7-13 (high) | 59 | 65 | 87 |
| B. October Survey: | | | |
| Score: | (*N*=30)<br>% | (*N*=37)<br>% | (*N*=132)<br>% |
| 0-5 (low) | (*3*) 60 | 41 | (*4*) 23 * |
| 6-10 (high) | 40 | 59 | 77 |

*Differences significant within the 5 per cent level of confidence or better include: 1 & 2; 3 & 4.

Although within the limits of the Pittsfield data, it is possible only to speculate as to the specific nature of the psychological processes relevant to attained level of political information, one might infer tentatively that lack of political knowledge indicates a general motivational and perceptual withdrawal on the part of the individual from at least the political phase of his community life. More probing research than was possible in the 1952 study of Pittsfield might throw light upon the underlying psychological processes.

The data presented in the following pages represent an attempt to specify and to measure certain psychological correlates of non-voting. Specifically, the following questions were raised: to what extent do we find within the chronic non-voting group a psychological identification with the three main political classifications,—Republican, Democrat, Independent? What is the nature of the pattern of assumptions (political perceptions) made by the members of the non-voter group regarding the political allegiance of various social groups? Finally, to what extent is non-voting correlated with level of no opinion on political issues?

If it is true that persistent failure to vote reflects psychological non-involvement in political affairs, one would expect to find within the non-voter group a comparatively low frequency of self-identification with recognized political divisions. In examining this possibility, the questions used in the Pittsfield surveys asked the respondents to indicate with which political group—Republican, Democrat, or Independent—they identified themselves. Our concern was not with the fact of membership as such, but rather, with what the individual considered himself to be politically. In Table 4, it will be seen

TABLE 4

POLITICAL GROUP IDENTIFICATION

| | Non-voter | Occasional voter | Voter |
|---|---|---|---|
| A. Frequency of Identification: | (*N*=126)<br>% | (*N*=143)<br>% | (*N*=665)<br>% |
| No political identification | (*1*) 28 | (*2*) 6 | (*3*) 3 * |
| Identification | 72 | 94 | 97 |
| B. Content of Identification: | (*N*=91)<br>% | (*N*=135)<br>% | (*N*=642)<br>% |
| Republican | 18 | 17 | 30 |
| Democrat | 29 | 31 | 29 |
| Independent | (*4*) 53 | (*5*) 52 | (*6*) 41 * |

*Differences significant within the 5 per cent level of confidence or better include: 1 & 2, 3; 2 & 1; 3 & 1; 4 & 6; 5 & 6; 6 & 4, 5.

that of the three respondent groups the non-voters most frequently professed no political identification. In addition to this finding, it should be noted that even though the proportion of non-voters claiming no political identification was significantly higher than was the case with the voter group, a substantial percentage, about 7 out of 10, of the members of the former group did identify with one of the three political classifications. Examination of the "B" part of Table 4, however, reveals that of this number, half called themselves "Independent." Almost identical results were found among the occasional voters. Among the individuals who consistently voted, however, this tendency was reversed with over half identifying with one of the two major political parties. These data show, therefore, not only a positive relationship between incidence of voting and incidence of self-identification politically, but also the frequency with which non-voters identify themselves as independents, a somewhat amorphous political classification.

In addition to being queried about the nature of their political group identifications, the Pittsfield respondents were asked a series of questions designed to measure their assumptions regarding the political allegiance of various social groups usually considered to be relatively stereotyped politically.[5] The assumption was made that if the non-voter was psychologically as well as behaviorally non-involved in the political affairs of his community, this fact would be reflected in the nature of his political perceptions. It was anticipated that the group of individuals who consistently exercised their voting privilege, in comparison to the non-voter group, would contain a larger percentage of respondents making a substantive assumption regarding the political behavior of various social groups, and further, that their assumptions would agree generally with the particular stereotype attached to the various groups being considered.

As indicated in Table 5, the latter expectation proved to be true with regard to only one of the five groups, whereas the former was borne out in the three cases. When asked how they thought the individuals composing the low economic group, and the low occupational category tended politically, the non-voter and the voter groups agreed in each case in terms of the frequency with which any answer at all was given. Among those individuals who did respond substantively, there was agreement as to the direction of the political allegiance attributed to the members of the social groups in question. These findings do not support our hypothesis.

The data obtained with reference to the other three groups, however, do support the proposed hypothesis. A significantly lower proportion of nonvoters professed any expectation regarding the political loyalties of people of high economic status, members of the Roman Catholic Church, and individuals classified occupationally as industrial executives. Further, in the latter case, the non-voters who did give substantive responses disagreed to a significant extent with their voter counterparts in the direction they expected the voting behavior of industrial executives to take.

The data included in Table 5, because it fails to support our hypothesis

---

[5] (1) "We would like to know whether, in your opinion, the following groups of Pittsfield people tend to vote Democratic, Republican, or Independent?
- (a) People earning less than $45.00 per week.
- (b) People living in Ward 4—for example, Pomeroy and Crofut Streets (an "A" residential area).
- (c) Members of the Roman Catholic Church.

(2) "What would be your best guess as to how people who work in factories are apt to vote—do you think . . . (interviewer supplies, one by one, each of the following job categories) will most likely vote for Eisenhower or Stevenson?
- (a) Laborers.
- (b) Foremen.
- (c) Office help.
- (d) Executives.

### TABLE 5
### ASSUMPTIONS REGARDING THE VOTING BEHAVIOR OF POLITICALLY STEREOTYPED SOCIAL GROUPS

| | Non-voter | | Occasional voter | | Voter | |
|---|---|---|---|---|---|---|
| A. Low Economic Group: | (*N*=49) % | | (*N*=46) % | | (*N*=247) % | |
| | "a" | "b" | "a" | "b" | "a" | "b" |
| Democrat | 47 | 66 | 46 | 70 | 59 | 76 |
| Republican | 18 | 26 | 15 | 23 | 9 | 12 |
| Independent | 6 | 8 | 4 | 7 | 10 | 12 |
| Don't know | 29 | | 35 | | 22 | |
| B. Low occupational level: | (*N*=30) % | | (*N*=37) % | | (*N*=132) % | |
| | "a" | "b" | "a" | "b" | "a" | "b" |
| Eisenhower | 13 | 17 | 14 | 16 | 14 | 16 |
| Stevenson | 67 | 83 | 70 | 84 | 74 | 84 |
| 50/50 split | 20 | | 16 | | 12 | |
| C. High Economic Group: | (*N*=49) % | | (*N*=46) % | | (*N*=247) % | |
| | "a" | "b" | "a" | "b" | "a" | "b" |
| Democrat | 6 | 12 | 2 | 4 | 3 | 4 |
| Republican | 45 | 88 | 54 | 92 | 77 | 94 |
| Independent | — | — | 2 | 4 | 2 | 2 |
| Don't know | (*1*) 49 | | 42 | | (*2*) 18 | * |
| D. Roman Catholics: | (*N*=49) % | | (*N*=46) % | | (*N*=247) % | |
| | "a" | "b" | "a" | "b" | "a" | "b" |
| Democrat | 41 | 74 | 48 | 78 | 50 | 70 |
| Republican | 8 | 14 | 9 | 15 | 6 | 9 |
| Independent | 6 | 12 | 4 | 7 | 15 | 21 |
| Don't know | (*3*) 45 | | 39 | | (*4*) 29 | * |
| E. High Occupational Level: | (*N*=30) % | | (*N*=37) % | | (*N*=132) % | |
| | "a" | "b" | "a" | "b" | "a" | "b" |
| Eisenhower | 40 | (*5*) 63 | 57 | 81 | 77 | (*6*) 89 * |
| Stevenson | 23 | 37 | 14 | 19 | 9 | 11 |
| 50/50 split | (*7*) 37 | | 29 | | (*8*) 14 | * |

*Differences significant within the 5 per cent level of confidence or better include: 1 & 2; 3 & 4; 5 & 6; 7 & 8.

consistently, might lead one to reject the proposed relationship between non-voting and perceptual or assumptive withdrawal. The writer feels, however, that it is more reasonable to infer that the inconsistency reflects the presence of additional relevant variables that were not controlled. For example, it is possible that the fact that the chronic non-voter group contains a large percentage of low economic and low occupational status individuals may be particularly germane to the obtained results. It is likely that even though, in a broad sense, an individual may be psychologically removed from the political affairs of his community, when asked to indicate the probable political behavior of the other members of his particular social groups,—groups with which he is likely to be identifying,—he is able to give a substantive answer.

The degree to which chronic non-voters profess no attitude or opinion regarding various current political issues was considered a further indication of psychological non-involvement. It was hypothesized that as level of non-voting increased so would the frequency with which no opinion was expressed on political issues. Although it was thought likely that the percentage of non-voters claiming to hold no opinion would vary from issue to issue, it was expected that in comparison to the other two groups of respondents,—the occasional voters, and the voters,—the non-voter group would consistently contain the highest percentage of people holding no opinion.

The data included in Table 6 represent a typical sample of the political

TABLE 6

FREQUENCY OF EXPRESSION OF NO OPINION REGARDING POLITICAL ISSUES

| | Non-voter | Occasional voter | Voter | |
|---|---|---|---|---|
| A. April Survey: | | | | |
| "In your opinion, what are the three most serious problems facing the United States today?" | | | | |
| | ($N$=49)<br>% | ($N$=46)<br>% | ($N$=247)<br>% | |
| Opinion | 82 | 83 | 94 | |
| No opinion | (*1*) 18 | (*2*) 17 | (*3*) 6 | * |
| B. August Survey: | | | | |
| "In your opinion, does organized labor—the CIO for example—have a great deal of influence in the politics of this country, or doesn't it have very much influence?" | | | | |
| | ($N$=32)<br>% | ($N$=32)<br>% | ($N$=145)<br>% | |
| Opinion | 53 | 72 | 92 | |
| No opinion | (*4*) 47 | (*5*) 28 | (*6*) 8 | * |
| C. October Survey: | | | | |
| "Do you think that Taft is helping or hurting Eisenhower's chances of being elected President?" | | | | |
| | ($N$=30)<br>% | ($N$=37)<br>% | ($N$=132)<br>% | |
| Opinion | 50 | 84 | 82 | |
| No opinion | (*7*) 50 | (*8*) 16 | (*9*) 18 | * |

*Differences significant within the 5 per cent level of confidence or better include: 1 & 3; 2 & 3; 3 & 1, 2; 4 & 6; 5 & 6; 6 & 4, 5; 7 & 8, 9.

issue questions asked in the April, August, and October Pittsfield surveys. It will be noted that not only did the expected relationship continue to hold true through the three survey periods, but also, that the relationship cross-cut a variety of issues differing widely in substance. While it would not be advisable to attempt to generalize from these results to the extent of assuming that the high degree of consistency in these data would necessarily hold true beyond the limits of this research effort, the Pittsfield findings do show a consistent and statistically significant trend.

A definitive study of the theme proposed in this article would involve considerably more penetrating research than was possible in the 1952 Pittsfield Project. Although the sample survey research technique permits one to investigate on a relatively wide scale certain variables relevant to the hypothesis that chronic non-voting reflects both subjective and behavioral withdrawal, it is clear that survey findings of the order presented above would assume greater meaning if reinforced by both clinical and experimental data. It is the writer's bias that field research is severely limited in terms of the hypotheses that one may realistically attempt to examine unless it is supplemented by controlled experimental analysis, and clinical study. In adopting a multi-methodological approach, it is more likely that one may be able to specify the characteristic motivational, cognitive, and perceptual patterns that constitute the core of overt behavioral phenomena such as chronic non-voting.

Recognizing the limits of the 1952 election study of Pittsfield, Massachusetts, the data obtained may be interpreted as supporting our central contention regarding the non-voter. The following variables were found to be related to low voting participation:

1. Low incidence of voluntary group membership.
2. Low exposure to political communication.
3. Low level of acquired political information.
4. Low frequency of political group identification.
5. Low frequency of substantive assumptions regarding the direction of the political behavior of various politically stereotyped social groups.
6. High frequency of expression of no opinion regarding current political issues.

## REFERENCES

1. CONNELLY, G. M., & FIELD, H. H. The non-voter—who he is, what he thinks. *Pub. Op. Quart.*, 1944, **8**, 175-187.
2. HASTINGS, P. K. Politics laboratory—1952. *J. Higher Educ.*, 1953, **24**, 12-17.
3. ———. The independent voter in 1952: A study of Pittsfield, Massachusetts. *Am. Pol. Sc. Rev.*, 1953, **47**, 805-810.

4. KEY, V. O., JR. Politics, Parties, and Pressure Groups. New York: Crowell, 1952.
5. SURVEY RESEARCH CENTER, UNIVERSITY OF MICHIGAN. A Study of the Presidential Vote: November 1948. Ann Arbor, 1949.
6. WOODWARD, J. L., & ROPER, E. Political activity of American citizens. *Am. Pol. Sc. Rev.*, 1950, **44**, 872-885.

*Department of Psychology*
*Williams College*
*Williamstown, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 313-319.

# A FURTHER STUDY OF ENTOPTIC STRAY LIGHT*

*Department of Psychology, Michigan State College*

S. Howard Bartley and Jean L. Seibel

## A. The Problem

The forms of stray light that usually have been considered to exist within the eye make up only a small part of all the light that is scattered there. There is a form of stray light that has been shown to be the basis for certain visual phenomena such as gamma movement and flicker in the visual field surrounding the test-target (2, 3). This stray light is the light reflected from the image of the target and cast onto the whole retina as a more or less uniform blanket. In addition to this, there is a source of stray light which seems to be attributable to scatter through the media themselves, thus forming a taper from the image down to the level of the general scatter just described.

Bartley (1) photometered the image-produced stray light and found the general blanket to be proportional to the intensity and area of the target (image) producing it. Bartley and Fry (4) measured the scatter from an intense image (glare source) detected through its effect in raising the differential threshold of a foveal target. Fry and Bartley (11) showed that the electroretinogram when a restricted target was imaged on the retina, was produced virtually entirely by the scattered light falling on the rest of the retina rather than by the light in the image. Riggs and Boynton (7) confirmed this recently in the case of the human electroretinogram with targets subtending visual angles from 1° to 12°. Later Boynton (5), using targets up to 100°, showed the relative contributions of the imaged and the non-focused or stray light in producing the electroretinogram. He also has directly photometered the stray light in several species of eye including the human (6).

Fry and Alpern (9, 10) verified the functional nature of stray light in the course of checking certain findings of Schouten and Ornstein (12) who had attributed certain phenomena to trans-retinal neural interaction.

---

*Received in the Editorial Office on June 17, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

The present paper reports an attempt by the flicker method to indicate the level of image-related stray light in relation to image area and intensity. It is known that critical flicker frequency (c.f.f.) of the target is related in an orderly way to target area and intensity. Since it is the target that is producing the stray light in question, there ought to be an orderly relation between target c.f.f. and field c.f.f. (i.e., c.f.f. at various points outside the target).

Virtually all, if not all measurements of target c.f.f. have involved targets of restricted area. In measurements here of field c.f.f. no such area restriction is involved. All the observer does is to report on flicker in a given local portion of the whole field at a given distance from the target. On this account, it was thought necessary to have at our command precise information on the relation between target area and target c.f.f. The literature did not provide us with this information in the way demanded, hence our first step was to obtain it for ourselves.

## B. Apparatus and Procedure

Intermittent light was delivered by a sectored disk whose rate of revolution was measured directly by a Weston tachometer. The light was interrupted at the nodal point of the lens system and cast onto an opal glass screen whose exposed portions formed, one by one, the target disks of the required areas. The general levels of light were produced by arrangements involving the inverse square law and filters. Minor variations in intensity were produced by manipulating lamp current. The senior author served as observer throughout.

The targets subtended visual angles from 51′ to 31°51′ as follows: A, 51′; B, 1°18′; C, 1°50′; D, 2°44′; E, 3°42′; F, 5°6′; and G, 31°51′. Area was measured in terms of the square of the visual angle, and the relative area range ran from 1 to about 1412.

## C. Results

Figure 1 indicates the findings on the relation between area and c.f.f. as intensity was varied. Each curve thus represents a separate area, the values of which were given in the previous section. It will be seen that the smaller areas produced curves without breaks as target intensity level dropped to the lower ranges, whereas the targets with angular subtenses of over 5° each manifested a definite hump. The lack of humps in the smaller targets is fortunate, for it works toward providing curves potentially more useful in comparing field c.f.f. findings with those of the targets.

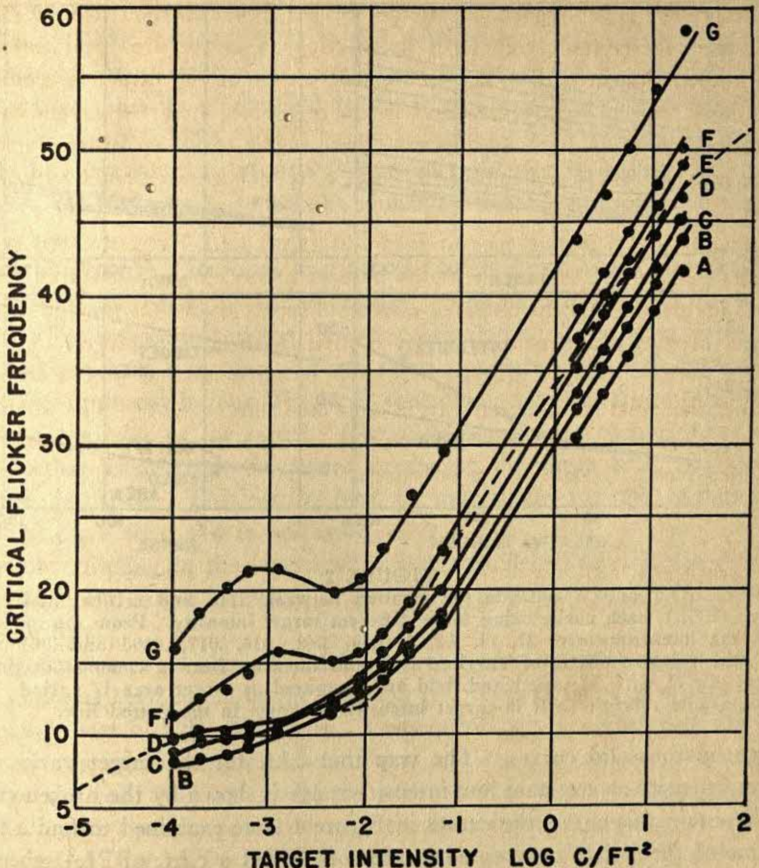


FIGURE 1

Curves showing the relation between critical flicker frequency (in cycles per second) and target intensity. Each curve is for a different target size. *A* is for a target subtending a visual angle of 51′; *B*, 1°18′; *C*, 1°50′; *D*, 2°44′; *E*, 3°42′; *F*, 5°6′; *G*, 31°51′. The light and no-light phases of the cycle are equal. Broken curve from Crozier and Wolf with a 14.3° target of moving stripes.

Figure 2 is composed of four parts. *A* and *B* show the relation between area and c.f.f., each curve representing a different target intensity level. In general, there is a linear relation between area and c.f.f. *C* shows the difference in the rate at which relative intensity and relative area affect c.f.f. as they are increased from a common starting point, the point at which c.f.f. is but 7.75 per second. In *D*, the c.f.f. for the target and for the field 6° from the target are compared. Both are linear as target area is increased, as is

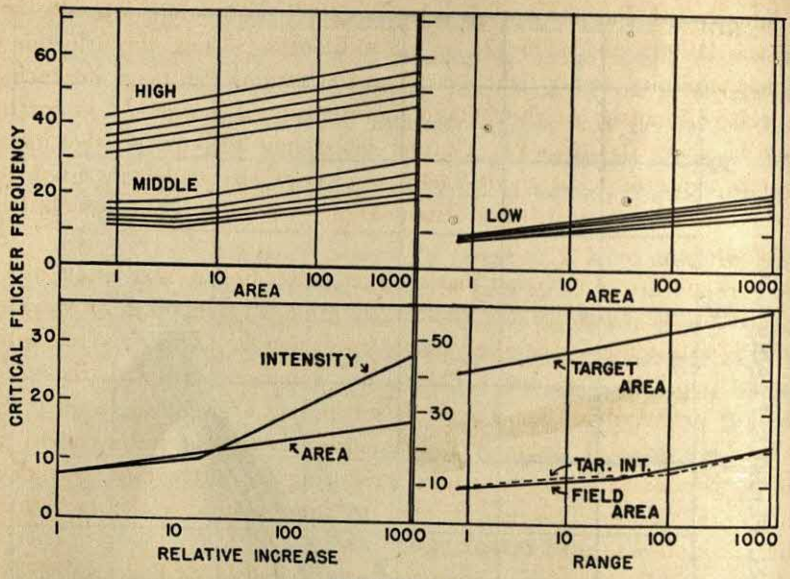


FIGURE 2

Curves in *A* and *B* indicate the relation between area and critical flicker frequency (c.f.f.), each curve being for a different target intensity. From top to bottom in *A*, the intensities are: 21, 11, 7.2, 3, 1.54, .068, .036, .017, .0097 and .005 c/ft.$^2$. In *C*, the relative effects of varying area and intensity from a common origin are shown. In *D*, c.f.f. of target and field are compared as target area is varied. Also a comparable relative shift in target intensity is shown in the dotted line.

shown by the solid curves. The way that c.f.f. for the target varies with target intensity at a certain low intensity range is shown by the broken curve.

To obtain this curve, the curves in Figure 1 were examined to find a point on one of them that represented a c.f.f. of 10 and a c.f.f. of 21.4 when the intensity level was increased over a range of three log units. Inspection of the curves in Figure 1 showed that on all three curves, *B*, *C*, and *D*, there was a point indicating a c.f.f. of 10; but that only on Curve *D* was it at an intensity value which if intensity were increased three log units, the c.f.f. would turn out to be in the proper neighborhood. Target *D*, in meeting these requirements, indicated a c.f.f. of 10 at an intensity of 0.0001 c/ft.$^2$. This means, then, that the c.f.f. curve for the field and the c.f.f. for Target *D* both began at 10 and ended at about 21.4 when they extended over an intensity range from 1 to 1,000. Knowing what the beginning intensity level was for Target *D*, we could infer that the field intensity must be the same. On this assumption, we could say that the field intensity with Target *A* (used for the field readings) producing a c.f.f. of 10 was 0.0001 c/ft.$^2$. Since the intensity of Target *A* itself was 21 c/ft.$^2$, the relative intensity of the field was 0.0001 divided by 21 or 4.8 millionths. Using the principle that the field intensity (stray light) varies as the area of the image producing it, the relative intensity of the field to the intensity of Target *F*, for example, would be about 0.0068. For a target subtending 90°, the relative intensity might be expected to be about 0.046, were pupil size to remain fixed under the consequent change in retinal illumination—which, of course, would hardly be the case.

Bartley and Fry (4) found that a target subtending a visual angle of 1°14′ at 5° from the *center* of the test-target, produced a blanket of stray light 0.000154 as intense as the glare source producing it. This occurred with an artificial pupil of 1 sq. mm. Their value compares well with the ratio of 0.000136 produced by our Target *B* subtending a visual angle of 1°18′ at 6° from the *edge* of the target. Obviously our present measurement point was further away from the source producing the effect than was that of Bartley and Fry. This would tend to compensate for the fact that no artificial pupil was used in our case.

Another finding in the investigation was that field c.f.f. decreased as distance from the target increased, or at least to a distance of 8°, if not somewhat further.

The question of which target c.f.f. curve to use in comparing target and field c.f.f.s has just been solved by the "cut and try" method of seeing which target-area curve matched the field curve. Consequently, it was Curve *D* that was selected. In support and corroboration of this choice is the fact when measurements of field c.f.f. are made, the c.f.f. reported upon is simply that of a small 2° or 3° patch concentric with the fixation point involving much the same area as Target *D*. This area always loses its flicker long before the flicker disappears in the broad experimentally undelimited field surrounding it. Thus it seems that although there is this broad field, it does not function as a unified whole as far as the moment of flicker disappearance is concerned.

It would seem further that in obtaining such field c.f.f.s, the principles ordinarily used in viewing targets are not used. With large targets, flicker does not disappear simultaneously all over their surfaces, but much sooner in some places than others. The proper c.f.f. reading is generally considered to be that representing the point at which all flicker has disappeared from within the whole target rather than simply in the central area concentric with the fixation point.

Another comparison of c.f.f.s with differing stimulus conditions is represented by the broken curve in Figure 1. This curve was reported by Crozier and Wolf (8) for a 14.3° field in which intermittency was produced by moving alternate black and white stripes. It will be seen that our small-area curves follow the Crozier-Wolf curve. Why this is so is not explained, but this comparison was resorted to as a check on the representativeness of the curves we obtained in our investigation. Actually, our results compared more closely to the Crozier-Wolf curve for their Observer *W*, than did Observer *C*'s curve in the same investigation compare with Observer *W*.

## D. SUMMARY

Using the flicker method, an attempt was made to determine the intensity of stray light produced at various distances from the retinal images of targets of various intensities and areas.

To provide material for comparing c.f.f. readings for target and surrounding field, c.f.f. curves for various target areas and intensities were obtained.

C.f.f.s for the field 6° from the target matched those obtained throughout a certain low intensity range for one area of target. From this the actual intensity of stray light on the retina was deduced.

Measurements of *relative* level of stray light to target area and intensity compared well with those previously obtained by Bartley and Fry when they studied the effect of stray light on raising the threshold of a test-object.

Field c.f.f varied in accord with distance from target edge for 8° or more.

## REFERENCES

1. BARTLEY, S. H. The comparative distribution of light in the stimulus on the retina. *J. Comp. Psychol.*, 1935, **19**, 149-154.
2. ———. Relation of stray light to flicker and the perception of movement. *Proc. Soc. Exper. Biol. & Med.*, 1935, **32**, 1180-1181.
3. ———. The basis of the flicker in the visual field surrounding the test-object. *J. Exper. Psychol.*, 1936, **19**, 342-350.
4. BARTLEY, S. H., & FRY, G. A. An indirect method of measuring stray light within the human eye. *J. Opt. Soc. Am.*, 1934, **24**, 342-347.
5. BOYNTON, R. M. Stray light and the human electroretinogram. *J. Opt. Soc. Am.*, 1953, **43**, 442-449.
6. ———. Physical measures of stray light in excised eyes. *Opt. Soc. Am. Program*, March, 1954.
7. BOYNTON, R., & RIGGS, L. A. The effect of stimulus area and intensity upon the human retinal response. *J. Exper. Psychol.*, 1951, **42**, 217-226.
8. CROZIER, W. J., & WOLF, E. Intensity and critical frequency for visual flicker. *J. Gen. Physiol.*, 1937, **21**, 203-221.
9. FRY, G. A., & ALPERN, M. The effect on foveal vision produced by a spot of light on the sclera near the margin of the retina. *J. Opt. Soc. Am.*, 1953, **43**, 187-188.

10. ———. The effect of a peripheral glare source upon the apparent brightness of an object. *J. Opt. Soc. Am.*, 1953, **43**, 189-195.
11. FRY, G. A., & BARTLEY, S. H. The effect of one border in the visual field upon the threshold of another. *Am. J. Physiol.*, 1935, **112**, 414-421.
12. ———. The relation of stray light in the eye to the retinal action potential. *Am. J. Physiol.*, 1935, **111**, 335-340.
13. SCHOUTEN, J. F., & ORNSTEIN, L. S. *J. Opt. Soc. Am.*, 1939, **29**, 168-182.

*Department of Psychology*
*Michigan State College*
*East Lansing, Michigan*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 321-329.

# MOTIVATION AND ATTENTION: A METHODOLOGICAL PROBLEM*

*U. S. Naval Training Center, San Diego, California*

---

FERDINAND STERN AND SAMUEL KARSON[1, 2]

---

A recent symposium on motivation has revealed a marked discrepancy in opinions as to the basic nature of this concept. The view is expressed that motivation is a distinctive process, a determinant of behavior, which is to be distinguished from an associative phenomenon (3). Stated more specifically this position holds that drives merely "activate or energize . . . (associative) tendencies" (3, p. 18). An opposing trend is evident in the belief that "motivational selectivity becomes a consequence rather than a direct determinant of such (associative) dispositions" (12, p. 58) although it is maintained that often this view "merely serves to push back the problem of motivational determinants" (12, p. 58). Other theorists emphasize that some manifestations of motivated behavior represent particular characteristics of man without stating unambiguously whether they consider motivation to be a distinctive process per se or merely an associative phenomenon (5, 10).

The central topic of this paper is the methodological problem of keeping psychological theory which aims at the explanation of events determined by the principle of causality free from all concepts alien to it. In contrast to natural science, which deals with the *is,* ethical systems, including the legal law, determine what the conduct of individuals *ought* to be. "The distinction between the 'ought' and the 'is' is fundamental for the description of (legal) law" (9, p. 37) as well as for the theoretical formulations of natural science. It seems, however, that this distinction is not always adhered to and that what appear to be fundamental concepts of psychology, such as motivation and attention, are actually ethical concepts; therefore, any attempts at incorporation of these concepts into psychological theory introduce logical fallacies which result in interminable controversy. It may well be that some theorists find it difficult to accept a strictly mechanistic determina-

---

*Received in the Editorial Office on June 17, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors gratefully acknowledge many helpful criticisms and suggestions by Dr. Virginia W. Voeks.

[2] The opinions or assertions contained in this article are the private ones of the writers and are not to be construed as official or reflecting the view of the Navy Department or the naval service at large.

tion of all behavior because it seemingly denies freedom of action and is therefore difficult to reconcile with the concept of responsibility which is a cornerstone of any social organization. Responsibility, however, which has been pointed out in a different context (16), is a concept of ethics and the legal law and not properly a concept of a natural science like psychology. The present authors hold that such a methodological error is to a great extent accountable for the distress experienced by many contemporary psychologists "at the thought that, unlike other scientists, they have not yet reached agreement on their basic repertoire of terms and concepts" (1, p. 138).

Similarly, the tendency in some traditional theories of jurisprudence is to ignore the borderline that separates a theory of legal norms which regulate human behavior from a science that causally explains actual human behavior. This failure results in a confusion of the problem of "how men legally ought to behave with the question as to how men actually do behave" (9, p. xv), resulting in the inclusion of concepts of natural science into legal theory. A brief digression is required to clarify this point. A few of the many pseudoproblems and fictions which arise from such a methodologically indefensible syncretism of methods will be illustrated: a statute is mistakenly considered the will or a command of the legislator; a law is held valid only in the case of conforming behavior by the individuals who are bound by this law; a dualism of state and law is created which raises the further problem of the applicability of the law to the state whose existence supposedly pre-dated the creation of the law, a problem that "plays such a great rôle, especially in German jurisprudence" (9, p. 197). Finally, one of the most paradoxical notions of all, that the state is an organism or at least "*organic* in its nature—meaning thereby, that its genesis and development, though partly artificial, are to a large extent . . . governed by natural laws" (15, p. 68).

The "pure theory of law" (9, p. xiii), a monistic theory formulated by Kelsen, "is directed at a structural analysis of positive law rather than at a psychological or economic explanation of its conditions, or a moral or political evaluation of its ends" (9, p. xiv). The theory demonstrates that the existence of a statute depends merely upon completion of the legislative procedure prescribed by the constitution and that an identification of the psychological processes involved in voting for a measure with willing its content is at best a fiction considering the frequent instances when members of the legislative assembly "have at most a very superficial knowledge of its contents" (9, p. 34). Kelsen's theory shows further that the validity of the legal order merely means "that men ought to behave as the legal norms prescribe" (9, p. 39); thus, validity is recognized as a "quality of law" (9, p. 40), and not of the actual behavior of the individuals for whom the law is binding, since this behavior may at times radically differ from the behavior demanded by the law. This theory eliminates the problem of the dualism of law and state as a theoretical problem by revealing the dualism to be a hypostatization of the "personification of a legal order" (9, p. 197). The concept of an organic nature of the state is recognized to be the outcome of a political ideology aimed at enhancing its prestige in the eyes of its citizenry.[3]

Returning now to motivation and attention, a critical analysis of the treatment of these problems may well reveal that they are pseudo-problems which result from a hypostatization of labels conveniently used to evaluate certain aspects of behavior. Many situations obtain in which only certain kinds of behavior are approved and other kinds of behavior are disapproved. Approval is sometimes expressed in stating that the subject has "motivation" or pays "attention" whereas disapproval is expressed by referring to lack of "motivation" or "attention." Failure to recognize that the terms motivation and attention are merely labels denoting value judgments (approval or disapproval) and that they are not determinants of behavior can lead to the notion that some behavior occurs without "motivation" or "attention." Now, once motivation and attention are not considered characteristics of *all* behavior, then they can easily acquire the status of determinants of *some* behavior and attempts at explaining their rôle leads to continuous controversy. It follows that any definitions of motivation or attention, whether conceived of as determinants of behavior or merely as associative phenomena, actually contain value judgments.

In his treatment of the so-called "drive for money" (3, p. 16) Brown distinguishes between drives which themselves are nondirectional, merely activating or energizing latent associative reaction tendencies and the specific reaction tendencies themselves. According to him unavoidable routine accidents during the first years of a child's life produce pain and anxiety. Concomitant display of worry by the parents establishes an associative cue for anxiety. If verbal complaints over the lack of money are included in the display of parental worry they also acquire a capacity for eliciting anxiety. Subsequent money-seeking responses are strongly reinforced due to reduction of anxiety following receipt of money. The drive for money then can be explained by "appeal to the cue stimuli attending anxiety and to the drive property of anxiety in activating the latent reaction tendencies attached to those cues" (3,

---

[3] In Kelsen's theory, all properties of the State are "presented as properties of a legal order" (9, p. 207): the territory of the State is "the territorial sphere of validity of the legal order" (9, p. 208) and its population is the "personal sphere of validity of this order" (9, p. 233).

p. 16). Brown's distinction between a nondirectional anxiety drive and directed associative reaction tendencies, however, does not necessarily follow from observation of actual behavior. From the behavioral point of view no pure nondirectional anxiety drive exists, but only anxiety reactions that are neither random nor nondirectional, but which are always jointly determined by the individual's past experiences, the present state of the organism, and the present combination of external and internal stimuli impinging upon him.[4] The contribution of past experience is singularly revealed by the *specific* direction of the money-seeking reactions. Relief of anxiety following receipt of money establishes a conditioned "goal-tendency" (4, p. 144). Thus a distinction between anxiety and subsequent money-seeking reactions which is based upon the criterion of direction suggests a value judgment concerning their effectiveness in relieving anxiety elicited by lack of money.

Harlow stresses the rôle of externally elicited motives without denying the importance of those motives that are internally aroused. He concludes that the "origin of the stimulating source is immaterial—internal or external" (5, p. 45) and that temporospatial contiguity of conditioned and unconditioned stimuli is the essential motivational condition for learning. His position, as Postman points out (12), amounts to stating that all stimuli are motivating agents and consequently, deprives motives and drives of their distinctive status. Harlow's view that the source of stimulation is immaterial holds only as long as there is no concern over what is being learned. As Hebb has pointed out, motivation is not a distinctive process and some learning is expected to occur even when the animal has not "first been deprived of food or injured by shock" (6, p. 181). Setting up a motivation merely means that behavior is channeled in the direction in which the experimenter is "interested" (6, p. 181). Hebb's position is of particular interest because of his avowed rejection of the stimulus-response point of view which provides the basic orientation of the present paper. Actually this agreement is not too surprising since association is also the basic principle of Hebb's theory. It should be mentioned, however, that his principle of sensori-sensory association (6, p. 70) differs from the *S-R* principle of association of conditioning theory which Hebb considers to be applicable only to "non-set-influenced" (6, p. 107) learning such as the "learned eyeblink" response (6, p. 108).

Harlow's paper (5) offers some hints as to why some theorists believe that *S-R* relationships are not sufficient for a complete explanation of certain characteristics of human behavior. One of these characteristics mentioned by him is that motivation may at times reach a high level of intensity and persistence; another, that this may occur even though such endeavors may prove useless or harmful.

---

[4] This problem is related to the clinical concept of free-floating anxiety, i.e., an anxiety reaction elicited by cues which the sufferer cannot identify.

Hutchinson's studies of sudden insight in creative thinkers (8) provide some important clues as to the nature of the motivational mechanisms suggested by such persistent efforts. He reports that sometimes the most strenuous attempts at solution of an intellectual task are stymied and lead to a temporary renunciation of the problem. This is followed at times by sudden insight which may occur in periods of relaxation or of preoccupation with apparently entirely unrelated problems. Hutchinson states "owing to the suddenness and apparent revelation of new material the experience is mistakenly looked upon by some as unrelated to past experience" (8, p. 324). This event can be explained by positing that the "irrelevant" situation contained a cue that elicited the associative link necessary for the continuation of the abandoned train of thought. Thus it appears that behavior at any given moment, which is the only basis for inferring varying degrees of motivation such as persistence or renunciation, is fully determined by *S-R* relationships. Similarly, Guthrie states that the essential condition for relinquishing a specific goal or a desire is "a *change in stimulation*" (4, p. 188).

Newcomb holds that the same principles of learning obtain for "social and non-social influences" (10, p. 145). He believes that learning is influenced by reward and punishment and motivated by anticipation or expectancy of rewards. The justification of motivation as an independent concept thus seems to depend on the rôle of expectancy. If expectancy is defined as an anticipatory goal response, one of the alternatives envisaged by Newcomb, then expectancy represents an associative phenomenon reducible to *S-R* relationships.

Postman's analysis of motivation in perception (13) suggests that apparent motivational selectivity is produced sometimes by "artifacts of performance" (13, p. 99) and further envisages the possibility of conceptualizing perceptual behavior as well as apparent motivational selectivity in terms of associative learning.

Hebb's treatment of attention supports the position that attention is not to be viewed as a separate process although this conclusion is somewhat obscured by his terminology. He holds that responses are "determined by something else besides the immediately preceding sensory stimulation" (6, p. 5) and that attention or perceptual set is the factor that influences the "selectivity of response" (6, p. 4). He conceives of attention in this sense as a "partly autonomous, or nonsensory, cerebral activity" (6, p. 11). A careful scrutiny of Hebb's various references to this factor of attention suggests that he attributes to it the rôle which in formulations of *S-R* theorists is assigned to associations learned in the past, an interpretation supported by his explicit statement that attention is "closely related to association" (6, p. 101). In the present context his rejection of unity of attention and his emphasis upon multiplicity of attention is of particular interest. The present authors hold that his position has to be extended since he limits it by stating that attention "may often be multiple" (6, p. 153) and that we "very often carry on two familiar activities at the same time" (6, p. 152). "Arguing and driving a car (is given as such an instance) . . . neither seems possible without 'attention'" (6, p. 152). Actual observation, however, suggests that the apparent division of attention in this case is the result of learning (practice), which is completed only after an individual has learned to respond to the various stimuli of the traffic automatically, i.e., without initiating his responses by his own verbal stimulations, a phase of driving usually characteristic of the beginner. This example of behavior with multiple attention can be interpreted as simultaneity of two activities which do not interfere with one another after practice and familiarity does not appear to be an indispensable requirement for simultaneity of two actions. Such an interpretation is supported by the results of Broadbent's experiment dealing with speaking and listening simultaneously (2). He found that interference between speaking and listening resulted in a significant decrease in the number of correct verbal answers to simple questions which sometimes were presented while the subjects were still answering a previous question. Broadbent does not offer a conclusive explanation for the reported observations. If we consider, however, that listening usually does not consist merely of auditory stimulation, but also includes speech of the listener, even though such speech will in routine situations be of a highly fragmentary and subvocal nature, then an interpretation of his results is suggested which implies the impossibility of conducting two conversations simultaneously. Two of Broadbent's specific findings add further support to this interpretation. A significant increase in the percentage of mistakes resulted from making the questions more concise through reduction of their length by omission of a stereotyped component which the subjects could easily learn to ignore. This modification deprived them of additional time for responding without interference from an immediately following question. Furthermore, the percentage of correct answers to questions which occasionally immediately followed another, increased significantly by a reduction of the length of the prescribed response without a concomitant shortening of the question. This improvement in performance may be attributable to the fact that the shorter response to the preceding question interfered less with listening to the immediately following question. Before closing the discussion of Hebb's treatment of multiple attention it must be stated that the "supposed . . . extraordinary unity about attention" (6, p. 151) criticized by him refers to meanings of attention as "state or end result of attending," or the "selectivity of the process," or "various properties of the mind which apparently cannot be defined" (6, p. 102). Since Hebb rejects these meanings of attention as "no longer precise enough" (6, p. 152) his theory is neither capable of proving or disproving multiplicity of attention if attention refers to one of these meanings. The concept of multiple attention following from his theoretical premises merely means that behavior at any given moment comprises several activities which are determined by an "immediately preceding sensory stimulation" (6, p. 5) and by a "partly autonomous, or non-sensory cerebral activity" (6, p. 11) which is the result of an associative process.

Careful observation of an active organism will reveal the impossibility of a complete analysis of its total behavior into discrete components and of identification of those components upon which attention supposedly is focused. Any such selection can be based only upon an arbitrary principle varying with different observers according to the importance ascribed to the aspect of behavior under observation. Berlyne holds that an "animal . . . will generally, unless it is asleep, tired or ill, be found to be doing several things at once but one thing in particular, i.e., it is expending much more energy on one group of responses than on others" and that "there is reason to believe . . . that if it were not concentrating on that activity it would be concentrating on something else" (1, p. 144). The apparent tendency to ascribe attention to important or the most important component of the total behavior, and the distinction between the healthy and the sick animal is apparently based upon an unrecognized value judgment. Further, the use of the criterion of output of energy as an indication of attention may well lead to surprising results in instances of strenuous physical work performed while preoccupied with one's thoughts. Hilgard's treatment of the problem of sustained attention suggests that his definition of attention as "perceptual focusing" (7, p. 295) does not refer to a particular quality of behavior at any given moment, but rather to the persistence of a certain course of action over an extended period of time. The impression of persistence, however, is merely an artifact of inaccurate observation because of the continuous changes of behavior. Thus the concept of persistence applied to behavior denotes a value judgment based upon the criterion of the appropriateness of a sequence of responses to accomplish a certain result.

Paschal believes that attention is an "act, not a state" (11, p. 402), an *"implicit anticipatory reaction"* (11, p. 397). He further holds that the "degrees of attention will correspond to the degrees of identity between the implicit anticipating and the subsequent overt reactions" (11, p. 397). This formulation of the problem of attention simply means that behavior at a given moment is influenced by its antecedents and that the existence of the pseudo-problem of attention is due to the fact that the antecedents sometimes include implicit components not readily observed. Guthrie states that attention "includes general movements of orientation, looking, listening, and so on, and includes also the inhibition of conflicting response" (4, p. 122). This description of attention is limited to an avowedly incomplete listing of some of its supposed components and does not provide any basis for an identification of attention per se. The impossibility of describing attention in strictly behavioral terms. adds further support to the position taken in this paper that attention merely refers to a value judgment of behavior.

The moral nature of the concept of attention is clearly recognizable in Ribot's theory of attention (14). He identifies voluntary attention with sustained effort which he believes did not exist prior to civilization. He contrasts the savages' passion for hunting, war, and gambling with work which he considers to be the "concrete, the most manifest form of attention" (14, p. 37). Many years ago Woodworth recognized the moral nature of the term attention. He states that " 'attention' is one of those nouns that are properly verbs" (17, p. 43) and that inattention means *"attention to something else"* (17, p. 50).

## SUMMARY

The goal of the present paper was to illustrate the inevitable creation of pseudo-problems which arise from a syncretism of methodologically incompatible concepts. Pseudo-problems necessarily will arise from the inclusion of value judgments into psychological theory. Precisely the reverse holds true for theories of law and state following an intrusion by the concepts of natural science into these realms. In support of this position reference was made to Kelsen's critical analysis of various legal concepts. The relevance of this position to psychological theory was shown by a discussion of motivation and attention which included recent formulations of the problem of motivation by Brown, Harlow, Newcomb, and Postman. The paper included a discussion of various theoretical formulations of attention and of multiple attention in particular. The results of this analysis seem to lend support to the position of the present paper that the terms *motivation* and *attention* refer only to apparently unrecognized value judgments.

## REFERENCES

1. BERLYNE, D. E. Attention, perception, and behavior theory. *Psychol. Rev.*, 1951, **58**, 137-146.
2. BROADBENT, D. E. Speaking and listening simultaneously. *J. Exp. Psychol.*, 1952, **43**, 267-273.
3. BROWN, J. S. Problems presented by the concept of acquired drives. In *Current Theory and Research in Motivation.* Lincoln: Univ. Nebraska Press, 1953.
4. GUTHRIE, E. R. The Psychology of Learning. New York: Harper, 1952.
5. HARLOW, H. F. Motivation as a factor in the acquisition of new responses. In *Current Theory and Research in Motivation.* Lincoln: Univ. Nebraska Press, 1953.
6. HEBB, D. O. The Organization of Behavior. New York: Wiley, 1949.
7. HILGARD, E. R. Introduction to Psychology. New York: Harcourt Brace, 1953.
8. HUTCHINSON, E. D. Varieties of insight in humans. *Psychiatry,* 1939, **2**, 323-332.
9. KELSEN, H. General Theory of Law and State. Cambridge: Harvard Univ. Press, 1949.
10. NEWCOMB, T. M. Motivation in social behavior. In *Current Theory and Research in Motivation.* Lincoln: Univ. Nebraska Press, 1953.
11. PASCHAL, F. C. The trends in theories of attention. *Psychol. Rev.*, 1941, **48**, 383-403.
12. POSTMAN, L. J. Comments on papers by Professors Brown and Harlow. In *Current Theory and Research in Motivation.* Lincoln: Univ. Nebraska Press, 1953.
13. ————. The experimental analysis of motivational factors in perception. In *Current Theory and Research in Motivation.* Lincoln: Univ. Nebraska Press, 1953.
14. RIBOT, T. The Psychology of Attention. Chicago: Open Court, 1898.
15. SMITH, G. H. The Theory of the State. Reprinted, Dec. 14, 1895, from the Proc. of the Amer. Philos. Soc., Vol. XXXIV.
16. STERN, F., & KARSON, S. A critique of Mowrer's theory of neurosis. (In preparation.)
17. WOODWORTH, R. S. Psychology. New York: Holt, 1945.

*4409 West Point Loma Blvd.*
*San Diego 7, California*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 331-338.

# FACILITIES OF THE EMORY UNIVERSITY PRIMATE BEHAVIOR LABORATORY*[1]

*Department of Psychology, Emory University*

ARTHUR J. RIOPELLE

## A. INTRODUCTION

This paper has a dual purpose. First, it describes the general structure, the animal cages, and the housekeeping facilities of the newly developed Primate Behavior Laboratory at Emory University. Secondly, it describes the principal test apparatus and the most frequently used test-trial procedures.

It is hoped that a description of the laboratory might prove useful in a number of ways to other persons who may be planning to engage in primate studies. Some of the problems encountered here are sure to be duplicated in other laboratories and our attempts to solve these problems may suggest other, more efficient, solutions. Also, it is hoped that suggestions for further improvements in this laboratory may be elicited from those who read this account. This description should also serve as a single reference source for future investigations of this laboratory which are conducted in the living cages and for studies in which the characteristics of the living room might be specific factors in the investigation. It also appears economical of printed space to have published separately a detailed description of the basic test apparatus and test-trial procedures so that individual research reports need only refer to this description.

## B. LABORATORY STRUCTURE

The Primate Behavior Laboratory occupies one large and three small rooms of a single-story frame building. The large room serves primarily as living quarters, whereas the three small rooms serve as testing rooms. The remainder of the building is occupied by a rat laboratory, a dark room, and a shop.

The animal room is 29½ feet long and 24½ feet wide. A schematic diagram of the layout of the animal room is presented in Figure 1. Along one

---

*Received in the Editorial Office on June 18, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The development of the laboratory, the construction of the equipment, and the publication of this report were supported in part by a research grant (M-589) of the National Institute of Mental Health of the National Institutes of Health, Public Health Service.

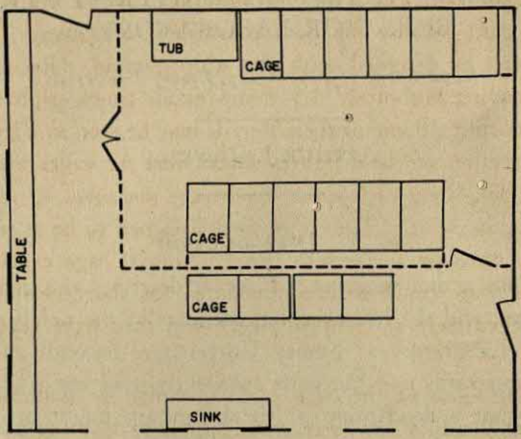


FIGURE 1
Schematic diagram of animal room. The dashed line represents the wire room. Six cages line the two opposing walls of the wire room. The sink and table are adjacent to the windowed walls.

windowed wall is a large double sink. Along the entire length of the other windowed wall is a table. At the opposite corner of this large room occurs a wire room that is 15 feet wide and 23 feet long. This wire room contains 12 living cages and a bathtub in addition to the daily needed cleaning equipment. The bathtub's usefulness has far exceeded our expectations for cleaning bulky cage bottoms and transport cages.

The animal room and the three test rooms are air conditioned. At this latitude the interior temperature of an uninsulated frame building would suffer daily fluctuations greater than would be desirable or healthful in the summer. Furthermore, the continuous flow of fresh air through the animal room makes the other laboratories habitable. In the winter these rooms are steam heated.

In addition to the daylight that enters through seven windows, the room may be artificially illuminated by fluorescent lights. Additional illumination may be obtained by turning on the four ultra-violet lamps which hang over the cages. These ultra-violet lamps are normally turned on for an hour or two each day. Each lamp requires 250 watts of power and services three cages. In addition to providing a source of vitamin *D*, radiation from these lamps tends to minimize skin eruptions and possibly to reduce laboratory odors. No claims are made for the effectiveness of these lamps, however, and the writer is willing to admit the possibility that the therapeutic effect may be human rather than simian.

The most significant aspect of the building was the wooden floor. This aspect is significant because the floor must therefore always be dry, and the living cages must be designed with this aim in mind. Recognizing the impossibility of having absolutely dry floors at all times, asphalt tile was laid on the wood flooring. Some of these details may be seen in Figures 2 and 3.

Further inspection of these figures shows that 12 cages are located inside the wire room, six along each side. Each cage measures 42 in. high by 43½ in deep by 27¾ in. wide. The cages were designed to be as wide as possible, yet would pass through a standard door. Normal cage capacity is two animals, although more could be accommodated for short periods. The living cages are made entirely of stainless steel,[2] and they have several features of importance:

1. The front door of the cage (side opening) is mounted on hinges in such a way that the door can be removed simply by opening and lifting. The door is easily cleaned in a tub of hot water. The large door covers the entire cage front and contains a guillotine door 18¼ in. high and 11¾ in. wide. Small transport cages modelled after those described by Brubach (1) and by Settlage and Harlow (4) may be used to pick up the animals. The transport cage may be placed flush against the living cage door before the guillotine door is raised. With this technique there is practically no risk of allowing an animal to escape.

2. A stainless steel perch eight in. wide extends across the rear of the cage at a height of 15 in. above the cage floor. This perch slopes down one-half in. toward the front so that liquid wastes will spill off the perch.

3. Eight inches below the perch is a watering cup. This cup (Johnson Poultry Cup, Cone Type, available through poultry supply houses) is permanently connected to the water mains and is fully automatic. The cup consists of an outer, covered, cup and an internal floating cup and operates in the following way. When an animal drinks from the cup he simultaneously drains the internal cup. This internal cup, when emptied, floats off a tiny inlet, thereby permitting water to enter both the outer and inner cups. Water ceases to enter the cups when the filled inner cup settles to the bottom. If desired, the outer cover of the watering cup may be removed so that continually running water can be obtained. (This feature is useful if the cage inhabitants are ill and reinfection is to be prevented). The reason the drink-

---

[2] The writer is indebted to many members of the Operations Department of Emory University for their coöperation and assistance in the construction of the cages and test apparatus. Major contributions to the design and construction of this equipment were made by Mr. L. C. Downey and Mr. T. G. Wilder.

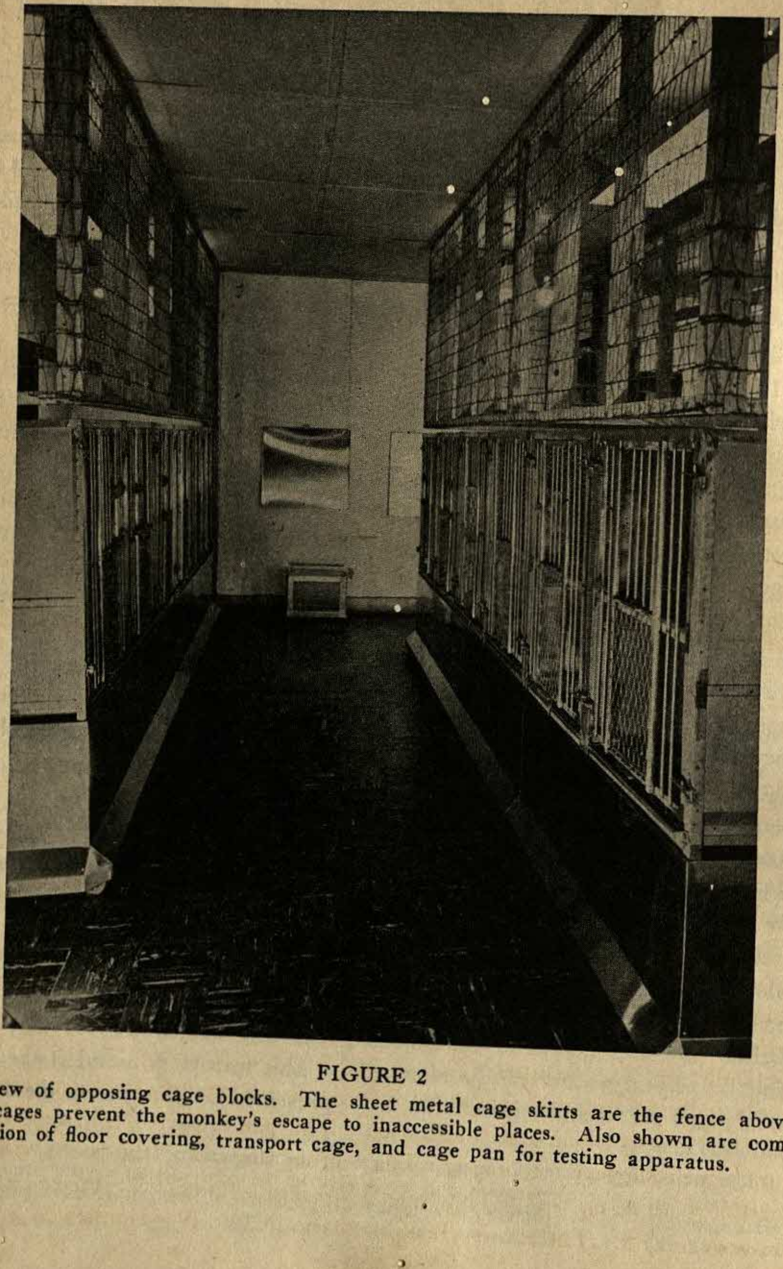


FIGURE 2

View of opposing cage blocks. The sheet metal cage skirts are the fence above the cages prevent the monkey's escape to inaccessible places. Also shown are composition of floor covering, transport cage, and cage pan for testing apparatus.

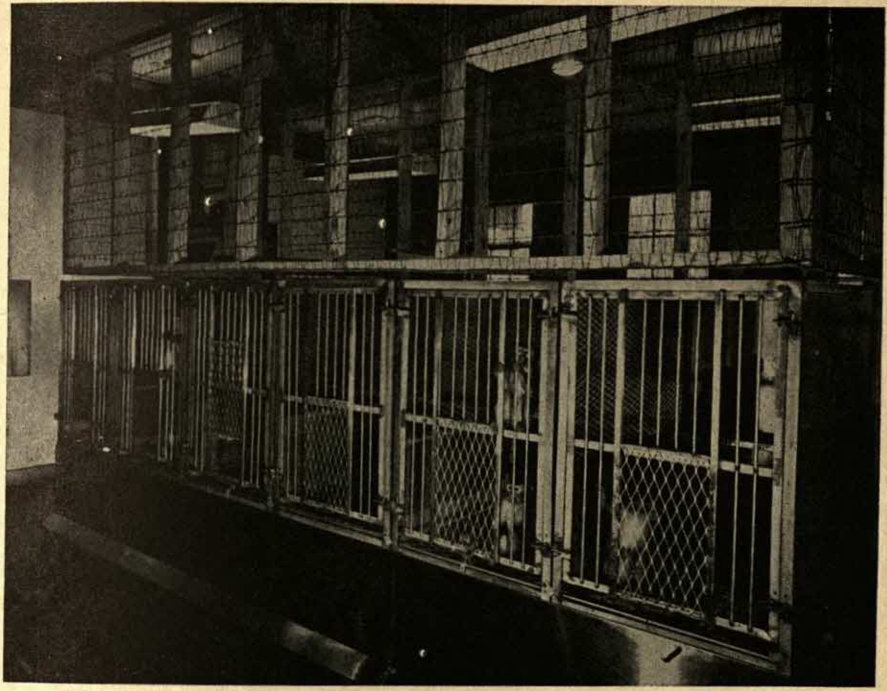


FIGURE 3
View of living cages. Details illustrated include long horizontal fluorescent lights, ultra-violet lamps, and construction of guillotine doors.

ing cup is located only eight inches below the perch is to discourage the animal's sitting on the cup.

4. The rear walls of the cages are of sheet metal so that the cages can be cleaned with a hose. Side walls between the cages are constructed of small mesh expanded metal. The removable floors and ceilings of the cages are constructed of large mesh expanded metal.

5. The space between the cage ceilings and the ceiling of the room is enclosed by a wire fence so that a free monkey can be easily captured. This wire barrier not only confines the free animal to a smaller space, it also protects the lighting facilities.

Six individual cages are mounted directly over a stainless steel pan approximately 14½ feet long and 4 feet wide. This pan is part of the same structure which supports the cages. The bottom surface of the pan slopes toward a spout at one end. This spout leads into a funnel connected to a four in. sewer drain. The pan is filled automatically and it is emptied by a flush valve. All connections to fresh water are made through rubber hoses so that the entire battery of cages, mounted on casters, can be rolled forward.

The schematic layout of the animal room, Figure 1, shows two racks of cages outside the wire room. Each rack contains two cages. These cages, not illustrated in the accompanying figures, are used as isolation cages for sick animals, for quarantine of new animals, and for spare animals. We have, for example, kept raccoons as well as other species of monkeys in these cages.

The three test rooms are located at the corner of the building opposite the animal room. They are roughly 9 x 12 feet in dimensions. The walls are treated with insulation to reduce sound transmission from one test room to the next. The rooms are air conditioned and equipped with a ceiling fan. The slight noise from these units serves also to mask unwanted noises.

## C. Basic Test Equipment

The primary test equipment used in this laboratory is a modified version of the Wisconsin General Test Apparatus (2). This instrument, when used for discrimination learning experiments embodies the principle of the Klüver form board (3). The Emory apparatus is pictured in Figure 4. As can be seen from this picture, the apparatus consists primarily of a barred restraining cage and an opposing table. A tray bearing the stimuli is mounted on casters and is pushed forward along the table top to the animal cage when the stimuli are presented. Each stimulus covers a food well.

For most experiments the animal is required to respond by displacing a stimulus from a food well to obtain food reward. Two counterbalanced screens are used to avoid artifacts. A one-way vision screen is interposed between the experimenter and the stimuli. When this screen is lowered, the experimenter can observe the actions of the monkey without being seen in return. An opaque screen lies immediately in front of the animal cage. This screen, when lowered, occludes the animal's vision of the stimulus tray. Attached to the bottom edge of the opaque screen and extending downward 12 inches from it is a quarter-inch thick sheet of transparent colorless plastic. The opaque screen may be lowered to a point where the animal may inspect the stimuli yet not displace one of them. Also, raising this opaque screen slowly from its lowest to its highest position encourages the animal to refrain from responding until it has had ample time to inspect all stimuli.

The detailed workings of the apparatus are best seen from a description of a complete test trial as employed in simultaneous discrimination. The

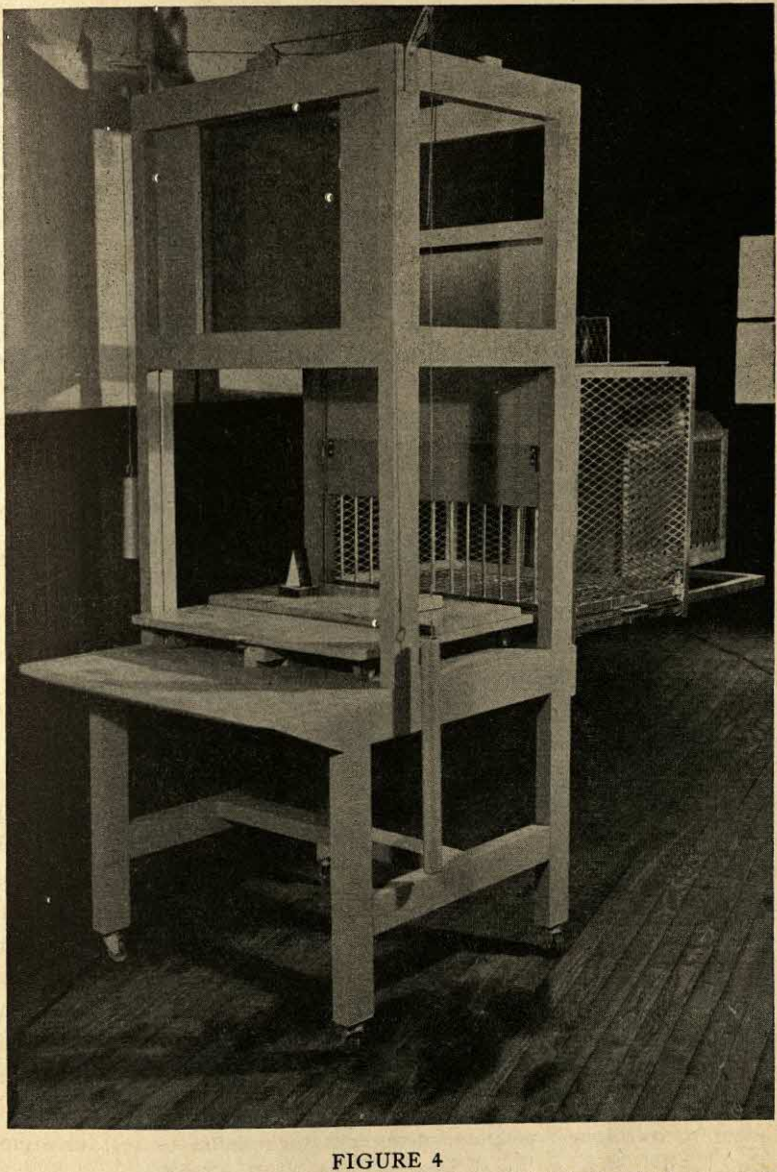


FIGURE 4

Basic test apparatus. To be noted are opaque screen (with plastic sheet attached), one-way vision screen, and sliding stimulus tray. The use of the transport cage is also illustrated.

trial begins with the opaque screen lowered. During this time the predetermined food well is baited and the stimuli are placed over the food wells simultaneously. The stimulus tray is then pushed forward, the one-way vision screen is lowered, and the opaque screen is raised slowly. The animal responds by displacing one stimulus. If correct, he is rewarded; if wrong, the stimulus tray is retracted. This sequence of events defines the non-correction procedure, in contrast to the situation wherein the animal, after committing an error, is allowed to displace the correct object to obtain food.

A variety of test foods have been employed with varying degrees of success, depending upon the particular animal and his previous test food experiences. Some foods which have been used are raisins, peanuts, grapes, diced orange, diced apple, and shelled corn. Standard practice in the laboratory is to maximize the variety within each test period.

## REFERENCES

1. Brubach, H. F. Improved cage designs for use in handling monkeys. *Science*, 1947, **106**, 18-19.
2. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51-65.
3. Klüver, H. Behavior Mechanisms in Monkeys. Chicago: Univ. Chicago Press, 1933.
4. Settlage, P., & Harlow, H. F. An effective and nontraumatic method of handling monkeys. *Science*, 1947, **106**, 300.

*Department of Psychology*
*Emory University*
*Emory University, Georgia*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 339-357.

# A FURTHER EXAMINATION OF THE RÔLE OF AUTISM IN A VISUAL FIGURE-GROUND RELATIONSHIP*

*The Menninger Foundation*

DOUGLAS N. JACKSON[1]

## A. THE PROBLEM

Ten years ago Schafer and Murphy (13) advanced the hypothesis that autism, defined as "the movement of cognitive processes in the direction of need satisfaction" (9, p. 365),[2] can serve as a determinant of figure-ground organization. To test this hypothesis, opposite complementary aspects of vertical contour lines, each resembling a human profile, were enclosed in semi-circles. These outline profiles were presented tachistoscopically to subjects and were either "rewarded" (subject was given money when the profile was shown) or "punished" (money was taken away). When rewarded and punished profiles were presented simultaneously so as to share the same contour line and form reversible patterns, each subject generally recognized previously rewarded profiles, but did not generally recognize previously punished material. Subjects were thus taught to perceive the ambiguous situation in one way rather than another. In other words, it was possible to control experimentally the process of *perceptual learning,* which the Gestaltists, Leeper (7), Tolman (15), Woodworth (19), have suggested may lie at the core of the conceptualization of the learning process. The experiment suggests the more general hypothesis that where figure-ground differentiation is possible, satisfying aspects of percepts tend to become organized as figure; the unsatisfying aspects being generally "excluded" and relegated to the ground. The experiment has been cited by those who consider the perceptual process not as an isolated sub-system, but capable of reciprocity at every stage with the state of the organism and all its complex needs.

---

*Received in the Editorial Office on June 19, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author gratefully acknowledges the many suggestions of Dr. Gardner Murphy in the design of the experiments and the preparation of the manuscript. Thanks are also due Drs. R. W. Gardner, R. R. Holt, P. S. Holzman, Irwin Miller, and Professor E. J. Asher for their careful reading of the manuscript.

[2] In current psychoanalytic usage the concept, "autism," generally refers to "primary process" phenomena where reality is grossly distorted or ignored. Bleuler, the originator of the concept, defined it more broadly, mentioning the "autism of waking normal people" but admitted that an "important gap" with respect to the intellectual aspect of "autistic thinking" existed in his presentation (1, p. 412).

Schafer and Murphy also found that after rewards and punishments were terminated, repeated stimulation led to figure-ground stability which may be independent of reward and punishment.

The experiment has been both criticized and defended. Pastore's (11) criticisms of the design and conclusions have been discussed elsewhere (4). Wallach (16), believing that figure-ground organization must precede memory trace arousal, predicted that the effects of reward and punishment would not be apparent if the outline figures were replaced by solid figures. This prediction was not confirmed in an experiment by Smith and Hochberg (14), who used electric shock punishment and solid black and white figures on a gray background. They found that punished aspects were significantly less frequently perceived in the post-training series, thus supporting the Schafer-Murphy hypothesis.

A repetition of the Schafer-Murphy experiment by Rock and Fleck (12) did not duplicate the original results. Since it is unlikely that discrepancies as large as those between the Schafer-Murphy experiment and the Rock and Fleck repetition occurred by chance and because the problem is theoretically important (2; 3; 9, pp. 362-390; 10), a systematic clarification was considered necessary. Two experiments were conducted[3] in an attempt to duplicate the Schafer-Murphy results and to determine whether the contradictory results of Rock and Fleck are due to weaknesses and previously undetected aspects of the original procedure, as Rock and Fleck suggested, or whether a procedural change introduced by the latter substantially altered the situation.

## B. Experiment 1

### 1. *Method*

The subjects for this experiment were 12 male students in introductory psychology classes from a midwestern university. The age range was 19 to 23, with a median of 20 years.

The Schafer and Murphy procedure was followed closely. Certain changes are noted below.

The experimental session was divided into a training and a post-training series. The training series consisted of 25 presentations of each of four profile outlines, totaling 100 presentations (Figure 1). The series was randomized with the exception that the same profile was not presented twice in succession.

---

[3] A third experiment substituting simple verbal "rewards" (e.g., "good," "that's right") for monetary influences was attempted, but results from six subjects suggested that these were not significantly effective in modifying figure-ground perception.

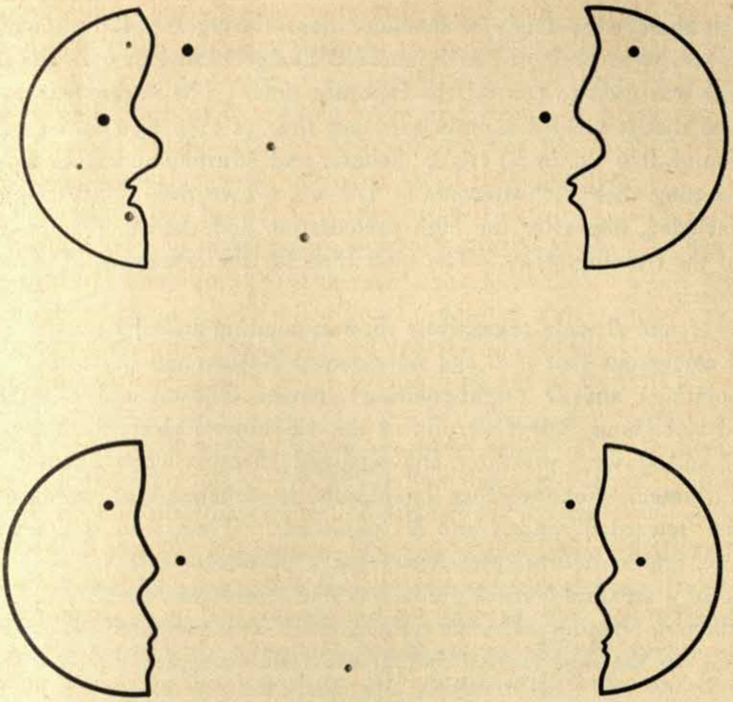


FIGURE 1
THE FOUR PROFILES PRESENTED IN THE TRAINING SERIES

The subjects, who knew in advance that they were to be paid 90 cents for participating, were seated in front of a modified Dodge tachistoscope, and six dimes and six nickels were placed on a table to their left. The following instructions were given:

> This is an experiment in vision. I am going to expose each of four profiles rapidly a number of times in this machine. When either of two of the profiles is shown, you will win 15 cents; when either of the other two is shown, you will lose 15 cents. You will not finish the experiment, however, with less than your 90 cents. Also, each profile has a name, and your task will be to learn the name that goes with each profile. To summarize what we shall do: (*a*) I shall announce the name of each face about to be presented. (*b*) The face will be presented in this machine. (*c*) You will repeat the name of the face. (*d*) I shall tell you whether you have won or lost 15 cents. (*e*) You will take or return 15 cents. Do you have any questions?

Since "rewarded" and "punished" faces were presented an equal number of times, all subjects finished the training series with their original 90 cents, although there were fairly wide fluctuations during the 100 presentations. The figures were presented in a modified Dodge tachistoscope. A mechanical timer was used to control the exposure time. The instrument was calibrated so that it yielded a mean exposure time of .336 sec. with a standard deviation of .026 sec. in 50 trials. Schafer and Murphy and Rock and Fleck report setting their tachistoscopes at 1/3 sec. Two five-minute rest periods were included, one after the 50th presentation and the other following the close of the training series. The total time for the experimental session was 50 minutes.

Faces *A* and *B* were respectively the left-pointing and right-pointing members of ambiguous Pair *A-B* and were named Nathan and Bertrand; Faces *C* (left-pointing) and *D* (right-pointing), named Clifford and Duncan, constituted ambiguous Pair *C-D*. Six of the 12 subjects received 15 cents each time *A* and *D* were presented and returned 15 cents when Faces *B* and *C* were presented. For the other six subjects the situation was reversed, *B* and *C* being "rewarded" and *A* and *D* "punished." Thus, since positional, idiosyncratic, and structural preferences for a particular face were controlled, the use of a separate control group was not considered necessary.

In the post-training series the training faces were presented simultaneously, so as to share the same vertical contour and form a complete circle (Figure 2).

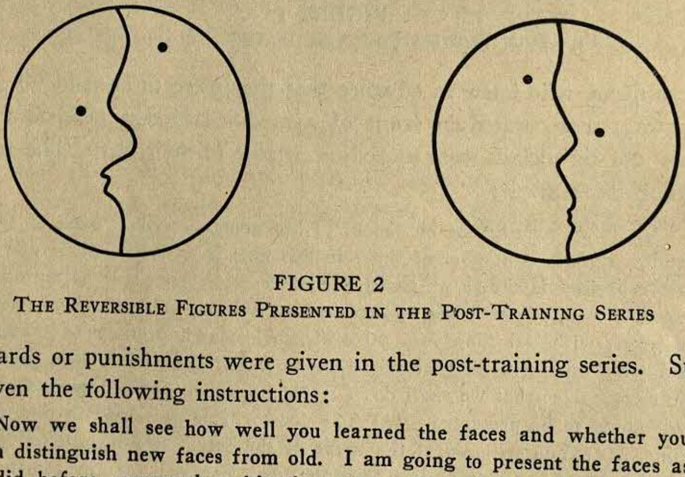


FIGURE 2
THE REVERSIBLE FIGURES PRESENTED IN THE POST-TRAINING SERIES

No rewards or punishments were given in the post-training series. Subjects were given the following instructions:

> Now we shall see how well you learned the faces and whether you can distinguish new faces from old. I am going to present the faces as I did before, except that this time they will be in a complete circle. All you have to do is to tell me the name of the face that you see. Quite frequently a new and strange face will be presented. When you do not recognize the face as one of those shown previously, you can indicate that the face is strange by saying "*X*." Do you have any questions?

The two new faces (Figure 3) were introduced as a precaution against the development of a directional set. Although they also contained a circle with a vertical profile line, they differed from Pairs *A-B* and *C-D* in that one profile (either the right-pointing or the left-pointing) was highly dominant.

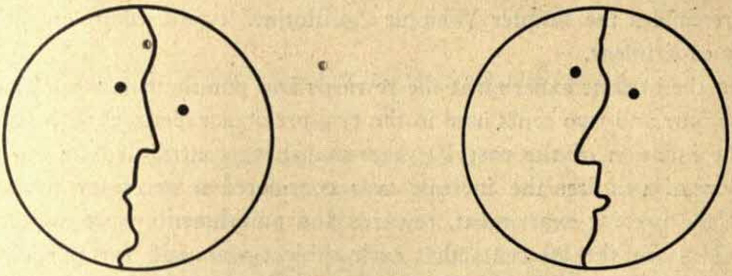


FIGURE 3
THE "SET-BREAKING" FIGURES PRESENTED IN THE POST-TRAINING SERIES

After each presentation of a training pair, the set-breaking profile facing in the direction opposite to subject's last response was presented. In the case of inappropriate responses to a training pair, e.g., an *"X"* or "Clifford" to the Nathan-Bertrand pair, it was not possible to determine whether a right or left-pointing aspect was perceived, and therefore the right and left set-breaking figures were alternated so as to equalize the number of presentations of each.

The post-training face-pairs were presented in randomized order, 16 times each, for a total of 32 presentations. Since each of these presentations was followed by a set-breaking figure, the total number of presentations for the post-training series was 63.

A post-experimental interview was conducted in which subjects were asked to describe the purpose of the experiment as they saw it, why they thought the money was included, what method they used to identify the faces, whether the faces seemed to have any particular expressions, and how they felt about winning and losing. An attempt was also made to determine whether they were aware of the reversibility of the patterns.

There are several differences between this experiment and those reported by Schafer and Murphy and by Rock and Fleck:

1. In the present experiment the diameters of the figures were 2½ in. and were presented at a distance of 24 in. in a modified Dodge tachistoscope. This approximated the conditions of the Schafer and Murphy experiment, in which a Whipple tachistoscope was used with figures with diameters of

2 to 3 in. presented at a distance of 18 in. It is in marked contrast to the Rock and Fleck repetition, in which a projector-type tachistoscope with images with diameters of 12 in. presented 54 in. from subject was used. The Dodge tachistoscope has a small, well defined, constantly lighted field, while the projector tachistoscope does not. The perceptual situation for the Dodge more closely resembles the simpler Whipple "guillotine" type used in the Schafer-Murphy experiment.

2. In the present experiment the rewards and punishments were increased from the four and two cents used in the two previous experiments to 15 cents. Since the inflation of the past 10 years may have contributed to a changed perception of pennies, the increase was considered a necessary precaution. Also, in the present experiment, rewards and punishments were added to or subtracted from the 90 cents that each subject was paid for participating, whereas in the previous experiments, subjects started with nothing and a small sum was allowed to accumulate.

3. In the two previous experiments subjects were asked to guess a number, the number ostensibly determining which training face was to appear and hence their winnings and losses. Subjects were told that the faces would be frequently renumbered and were led to believe that some "system" was in operation. Since a predetermined randomized order had been prepared, this was not actually the case. To be sure that perception was actually being rewarded and punished and to eliminate deception, a direct rewarding and punishing of the percepts, arbitrarily presented, was substituted for the "number guessing" procedure.

4. A new face pair (Figure 2, second pair) was substituted for one of those used in the previous experiments (13, p. 337, Figure 2, second pair). This was found necessary because preliminary experimentation (without monetary influences) indicated that recognitions of the left-pointing member of the original pair were disproportionate.

5. In the two previous experiments a strange profile was presented late in the training series. The subject was told that this was to insure his paying attention and that if he failed to recognize this "phony," he would be penalized 10 cents. Since this device was considered to be of dubious relevance, it was eliminated from the procedure.

## 2. *Results*

Table 1 presents a summary of the responses of each subject to the rewarded and punished figures.

It will be noted that for the first 16 presentations of the ambiguous material

### TABLE 1

Number of Responses to Post-Training Series Test Figures and Set-Breaking Figures: Experiment 1
($N$ for each subject = 63; total $N$ = 756)

| | | Responses to *A-B* and *C-D* test figures | | | | | | | | | Responses to *X* figures | | Total number of | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | First 16 trials | | | Second 16 trials | | | Entire series | | | Entire series | | | |
| Condition | Subject | R | P | M | R | P | M | R | P | M | X | M | Correct | Incorrect |
| A and D rewarded, B and C punished | 1 | 11 | 1 | 4 | 13 | 2 | 1 | 24 | 3 | 5 | 25 | 6 | 52 | 11 |
| | 2 | 6 | 1 | 9 | 16 | 0 | 0 | 22 | 1 | 9 | 31 | 0 | 54 | 9 |
| | 3 | 15 | 0 | 1 | 16 | 0 | 0 | 31 | 0 | 1 | 21 | 10 | 52 | 11 |
| | 4 | 9 | 7 | 0 | 13 | 3 | 0 | 22 | 10 | 0 | 29 | 2 | 61 | 2 |
| | 5 | 4 | 4 | 8 | 4 | 9 | 3 | 8 | 13 | 11 | 31 | 0 | 52 | 11 |
| | 6 | 5 | 10 | 1 | 3 | 12 | 1 | 8 | 22 | 2 | 31 | 0 | 61 | 2 |
| Total, subjects | 1-6 | 50 | 23 | 23 | 65 | 26 | 5 | 115 | 49 | 28 | 168 | 18 | 332 | 46 |
| B and C rewarded, A and D punished | 7 | 13 | 0 | 3 | 14 | 0 | 2 | 27 | 0 | 5 | 31 | 0 | 58 | 5 |
| | 8 | 4 | 2 | 10 | 3 | 4 | 9 | 7 | 6 | 19 | 30 | 1 | 43 | 20 |
| | 9 | 3 | 5 | 8 | 2 | 5 | 9 | 5 | 10 | 17 | 25 | 6 | 40 | 23 |
| | 10 | 14 | 1 | 1 | 16 | 0 | 0 | 30 | 1 | 1 | 15 | 16 | 46 | 17 |
| | 11 | 11 | 4 | 1 | 6 | 10 | 0 | 17 | 14 | 1 | 30 | 1 | 61 | 2 |
| | 12 | 10 | 4 | 2 | 10 | 6 | 0 | 20 | 10 | 2 | 8 | 23 | 38 | 25 |
| Total, subjects | 7-12 | 55 | 16 | 25 | 51 | 25 | 20 | 106 | 41 | 45 | 139 | 47 | 286 | 92 |
| Total | | 105 | 39 | 48 | 116 | 51 | 25 | 221 | 90 | 73 | 307 | 65 | 618 | 138 |
| Chi-square (R, P; 1 *df*) | | | 15.01 | | | 12.42 | | | 31.31 | | | | | |
| *p* | | | <.001 | | | <.001 | | | <.001 | | | | | |

*Notes* The meanings of the symbols used are as follows:

*R*—Recognitions of the previously rewarded figure in reversible pattern.
*P*—Recognitions of the previously punished figure in reversible pattern.
*M*—Mistaken identification of figures.
*X*—Responses to set-breaking figures.

nine of the 12 subjects predominantly recognized the previously rewarded face, one with an equal number of recognitions of previously rewarded and punished material, and two predominantly recognized previously punished faces. For the entire series of 32 presentations, nine responded in the direction of previous reward and three in the previously punished direction. All of Schafer's five subjects responded predominantly in the direction of previous reward, while Rock and Fleck report that their six subjects were divided equally between responding in the direction of rewarded and in the direction of punished material.

When the frequency of recognition of previously rewarded and punished faces from chance expectancy were compared, highly significant differences were found. Three different methods, critical ratio, chi-square, and binomial expansion, all yield differences between recognitions of rewarded and punished profiles well below the .001 significance level for the post-training series. The more stringent test of individual susceptibility to reward and punishment for the first 16 trials yields a distribution in which nine subjects responded in the rewarded direction, two in the punished, and one responding equally to reward and punishment. Disregarding the latter subject, the binomial distribution yields a $p < .02$ with the one-tail test when individuals are compared with respect to susceptibility to reward and punishment.

Table 2 presents the responses of each subject to each face in the ambiguous situations. It indicates, contrary to the findings of Rock and Fleck, that

TABLE 2
NUMBER OF RESPONSES TO EACH ASPECT OF POST-TRAINING SERIES AMBIGUOUS TEST FIGURES *A-B* AND *C-D*: EXPERIMENT 1

| Condition | Subject | First 16 trials | | | | Entire series | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | A | B | C | D | A | B | C | D |
| A and D rewarded, B and C punished, | 1 | 6 | 0 | 1 | 5 | 14 | 0 | 3 | 10 |
| | 2 | 1 | 1 | 0 | 5 | 9 | 1 | 0 | 13 |
| | 3 | 7 | 0 | 0 | 8 | 15 | 0 | 0 | 16 |
| | 4 | 5 | 3 | 4 | 4 | 10 | 6 | 4 | 12 |
| | 5 | 1 | 3 | 1 | 3 | 5 | 5 | 8 | 3 |
| | 6 | 5 | 3 | 7 | 0 | 7 | 8 | 14 | 1 |
| Total, subjects | 1-6 | 25 | 10 | 13 | 25 | 60 | 20 | 29 | 55 |
| B and C rewarded, A and D punished | 7 | 0 | 7 | 6 | 0 | 0 | 15 | 12 | 0 |
| | 8 | 2 | 2 | 2 | 0 | 6 | 5 | 2 | 0 |
| | 9 | 1 | 0 | 3 | 4 | 1 | 0 | 5 | 9 |
| | 10 | 0 | 8 | 6 | 1 | 0 | 16 | 14 | 1 |
| | 11 | 2 | 5 | 6 | 2 | 10 | 5 | 12 | 4 |
| | 12 | 1 | 7 | 3 | 3 | 1 | 15 | 5 | 9 |
| Total, subjects | 7-12 | 6 | 29 | 26 | 10 | 18 | 56 | 50 | 23 |
| Total | | 31 | 39 | 39 | 35 | 78 | 76 | 79 | 78 |

results from both stimulus pairs are in the rewarded direction. Also, the totals afford no evidence for directional or positional preferences.

To indicate that the effects of reward and punishment on figure-ground perception is consistent for the two ambiguous pairs, recognitions of previously rewarded profiles in one pair were correlated with recognitions of previously rewarded profiles in the other pair. The same was done with recognitions of previously punished profiles. By correlating, for example, the number of recognitions of Face *A* with recognitions of Face *D* for Subjects 1 to 6, and recognitions of Face *B* with recognitions of Face *C* for subjects 7 to 12 (all of which were previously rewarded for the subjects mentioned), structural, idiosyncratic, and positional preferences are systematically randomized. If the latter were the sole source of variance, zero correlations would be expected. Table 3 indicates that this is not the case; subjects tend to respond consistently and reliably to the two pairs. This does not agree with the findings of Rock and Fleck, who state: "In examining the individual results for both situations [profile pairs] combined, it appears that, although the total number of rewarded and punished faces perceived is about equal, this total derives from highly *unequal* distributions [of each reversible pair] for each subject" (12, p. 771, italics in original).

TABLE 3

INDIVIDUAL CONSISTENCIES IN SUSCEPTIBILITY TO REWARD AND PUNISHMENT—CORRELATIONS TO COMPARE AMBIGUOUS TEST FIGURES *A-B* AND *C-D* WITH RESPECT TO RECOGNITIONS OF REWARDED AND OF PUNISHED ASPECTS: EXPERIMENT 1 ($N = 12$)

| | First 16 trials | Second 16 trials | Entire series |
|---|---|---|---|
| Correlation between rewarded aspect of A-B with rewarded aspect of C-D | .67** | .50* | .57* |
| Correlation between punished aspect of A-B with punished aspect of C-D | .59* | .14 | .37 |

*Significant at or below the 5% point (one-tail test).
**Significant at or below the 1% point (one-tail test).

It is noteworthy that in the second 16 presentations there were significantly more correct identifications ($X^2 = 6.60$, 1 *df*, $p < .02$) than in the first 16 presentations. The post-training series was a new situation with more difficult figures. This may have caused some disruption of the previous learning requiring habituation and new perceptual learning. This same process may have been responsible for the "consolidation" effect reported by Schafer and Murphy. The authors suggest that after the 16th trial, the effects of reward and punishment diminished and recognitions became consolidated so that one aspect of the pair was consistently seen as figure. Although the total effects of the reward and punishment were not as great in the present experiment after the 16th trial, they remained, nevertheless, highly significant. We cannot isolate in the present experiment, as did Schafer and Murphy, the effects of sheer repetition and the effects of reward and punishment. The test of the consolidation hypothesis awaits an experiment where this variable may be independently studied.

The Rock and Fleck data indicate a considerable number of incorrect identifications of the training and the set-breaking faces, many more than reported by Schafer and Murphy. The authors, in the discussion of their results, attach considerable importance to this finding, suggesting that because of the guesswork and uncertainty in the post-training series, any significant results obtained may be a function of this rather than previous reward and punishment. Approximately one-third of the responses of their experimental group and over half of the responses of their control group were incorrect identifications. The subjects in the present experiment had significantly more success in identifying the figures correctly (Table 1) than the Rock and Fleck experimental subjects ($X^2 = 25.5$, 1 *df*, $p < .001$). On the other hand, when correct and incorrect recognitions from the first 16 trials of the post-training series of the Schafer-Murphy experiment are compared with the same data from the present experiment, the difference is not significant ($X^2 = 2.0$, 1 *df*, $p > .10$). Thus, unlike the Rock and Fleck experiment, the present experiment yields comparable results to those of Schafer and Murphy with respect to accuracy of recognition of post-training faces.

### 3. *The Post-Experimental Interview*

The subjects were questioned about the purpose of the experiment. Most subjects gave such answers as "testing your recognition," or "testing perception and retention." However, one subject (No. 1) correctly replied that the purpose was "probably to get an association of winning 15 cents with Nathan and Duncan." The subjects were also asked how they felt about winning and losing the money. Like the Rock and Fleck subjects, in every case their answer to this question was that they were either indifferent to the money or were distracted by it. Since it was clear that the rewards and punishments were having some effect upon their perception, further "probing" was considered necessary. This included "leading" questions, so as to give subjects ample opportunity to admit any feelings they may have had about the money. Replies to these questions may be roughly classified into two categories: (*a*) Those who continued to deny any feelings about the money (Subjects No. 5, 6, 8, 9, 11); and (*b*) those who, after first denying any feelings, indicated in some way that they were concerned about the money.[4] The following sequence is typical of the latter group: (Subject No. 2). "How did you feel about winning and losing?" "I didn't pay any attention to it." (Further questioning.) "Except when it got down pretty low I thought it better take a turn for the better soon." Thus, although the subjects' replies may represent a genuine lack of interest in the money in the experimental situation, in many cases they cannot be taken at face value The "need" for money may be a culturally ingrained pattern, but it may be equally important for an individual to deny concern over trivial sums of money "won" or "lost" in an artificial and purely experimental situation in which nothing much was at stake and for which he had volunteered.

In reviewing the results of subjects who indicated concern over rewards and those who did not, it is clear that the subjects who continued to deny any feelings about the money were the ones who showed the strongest tendency to recognize previously punished faces. For the total series, subjects who admitted feelings about the money recognized 176 rewarded profiles and 25 punished ones, while those who continued to deny feelings recognized only 45 rewarded profiles and 65 punished ones. This difference is significant ($X^2 = 72.99$, 1 *df*, $p < .001$). Also the subjects who continued to deny feelings about the money made significantly more mistakes in recognizing test figures than subjects who admitted feelings ($X^2 = 25.34$, 1 *df*, $p < .001$). These differences demonstrate that the subjects who failed to show the effects of reward and punishment (as assumed by the experimenter) actually do not run counter to the hypothesis when their own motivation is considered. This helps to avoid the problem, so often encountered in motivation research, of the experimenter's confusing his interpretation of what the subjects' motivation should be with what it actually is. Moreover, these differences in motivation and ability to learn, as crudely defined in this experiment, may yield clues to basic individual differences.

The interview further revealed that seven of the subjects perceived the profiles as if they were those of actual persons [even in the manner which Werner (18) calls physiognomic], reporting that the faces had particular expressions. There was a tendency, although not entirely consistent, to perceive rewarded faces as having more desirable expressions, e.g., "angelic," or "happy," and punished faces as having less socially acceptable expressions, e.g., "ape-like," or "stupid looking."

---

[4] Four psychologists judged the interview protocols as a check upon the experimenter's classification. The judges agreed with the experimenter to the extent of 83 per cent and their average rating did not affect the statistical significance of differences between the groups.

## C. Experiment 2

### 1. *Procedure*

The procedure for Experiment 2 was similar to the one outlined above in all essential respects, with the following exceptions:

1. A Keystone Slide Projector was used with 300-watt illumination. This cast an image with an 8½-in. diameter upon a white screen 69 in. from the subject. Rock and Fleck also used a projector tachistoscope, but it was an opaque projector rather than a slide projector.[5]

2. In addition to the profile pairs used in the earlier experiments, four additional pairs were added. This necessitated using three groups of six subjects for each of the two profile-pair combinations, making a total of 18. It was necessary to work with a total of 23 subjects to get 18 who could meet the criterion of six recognitions in the first 16 presentations, since some subjects had difficulty in recognizing the faces in the post-training series of reversible figures. It was assumed that those who made fewer than six recognitions had not learned the profiles sufficiently. The 18 subjects finally used were composed of 14 students from the same university as those who participated in Experiment 1, two Air Force officers, and two Menninger Foundation secretaries. The age range was 19 to 41 with a median of 21 years. All but three subjects were in the age range of 19-29.

### 2. *Results*

The results from Experiment 2 are in sharp contrast with those reported above. Table 4 presents data for the 18 subjects under the different conditions. The profile pairs presented to Group II (Nos. 3 and 4) were the same as those used by Schafer and Murphy and by Rock and Fleck, and pair No. 5 is the same as Pair *C-D* described above.

The experiment was originally designed to test the effects of reward and punishment, determining the contributions of the different reversible pairs and the extent of subject variations by analysis of variance. However, many of the same difficulties reported by Rock and Fleck were encountered here. Prominent among these were: (*a*) many mistakes in recognizing the reversible patterns; and (*b*) highly unequal distributions for the same subject in recognitions of rewarded and punished profiles in the different reversible situations.

---

[5] Dr. Rock supplied this information in a personal communication.

Table 4 indicates that although the over-all effect of reward and punishment is statistically significant, this tendency is not nearly so clear as in Experiment 1. When the effects of rewards and punishments are considered individually for the first 16 presentations and for the second 16 presentations, neither one is significant at the five per cent level as a result of the smaller number. Also, for the first 16 presentations, the effects of reward and punishment in Experiment 1 are significantly greater than those in Experiment 2 ($X^2 = 7.98$, 1 *df*, $p < .01$). The number of incorrect recognitions of profiles in Experiment 2, like the comparable Rock and Fleck data, is significantly greater than those in Experiment 1 ($X^2 = 17.52$, 1 *df*, $p < .001$).

TABLE 4

NUMBER OF RESPONSES TO SIX AMBIGUOUS POST-TRAINING FIGURES: EXPERIMENT 2*

($N$ for each subject = 32; total $N$ = 576)

| Group | | Subjects | Figure | Condition | R | P | M |
|---|---|---|---|---|---|---|---|
| I | a | 3 | 1 and 2 | A and D+, B and C— | 30 | 39 | 27 |
| | b | 3 | | B and C+, A and D— | 52 | 31 | 13 |
| II | a | 3 | 3 and 4 | A and D+, B and C— | 48 | 24 | 24 |
| | b | 3 | | B and C+, A and D— | 35 | 29 | 32 |
| III | a | 3 | 5 and 6 | A and D+, B and C— | 36 | 19 | 41 |
| | b | 3 | | B and C+, A and D— | 30 | 22 | 44 |
| Totals | | 18 | | | 231 | 164 | 181 |

Chi-square (R, P) = 5.39 (1 *df*), p <.05.

*Note.* + indicates previously rewarded profiles.
— indicates previously punished profiles.

*A more comprehensive table, giving the number of recognitions of each of the training figures, as well as the "*X*" figures, for each subject may be secured by ordering Document No. 4295 from American Documentation Institute, 1719 N. Street, N. W., Washington 6, D. C., remitting $1.25 for 35 mm. microfilm or $1.25 for a 6 by 8 in. photocopy.

Confronted with the difficult task of recognizing projected images in a new situation, the subject is forced to seek some kind of "anchorage," some mode of perceiving the ambiguous situation so that his probability of correct identifications will increase. One way of doing this is by consistently looking to a particular side for the profile and responding in terms of a stereotyped position set. The results for Experiment 2 indicate that the highly unequal distributions for most subjects of the effect of reward and punishment on the two pairs is a function of a pervasive predisposition to select left-pointing faces. For the 32 presentations the totals are 305 recognitions of left-pointing faces and 90 of right-pointing faces. This difference is statistically significant ($X^2 = 98.29$, 1 *df*, $p < .001$). The effect was present in every reversible pair. At the close of the session 11 subjects in the experimental group were asked to draw profiles and *all* drew them pointing left. Subjects' preference for left-pointing figures in this situation may reflect a general way of perceiving.

A recent study by Mishkin and Forgays (8) indicates that tachistoscopic recognition of words is greater in certain parts of the right visual field than in corresponding parts of the left visual field. The authors' conclusion, that selective retinal training arising from the reading situation accounts for the development of cell assemblies to facilitate right peripheral vision, may be appropriate in relation to the present results.

It is important to note that in the "easier" situation reported in Experiment 1, the totals for responses to right-pointing and to left-pointing faces are practically identical (right-pointing, 157; left-pointing, 154). Rock and Fleck noted the highly unequal distributions for each subject for the two reversible pairs but did not relate this to positional sets. When their data for the experimental group are re-examined with respect to left-pointing face dominance, they show a tendency in the direction of left-pointing predominance, which, although approaching statistical significance, must be interpreted with caution because of the small number of subjects and the individual inconsistencies. This position dominance, very pronounced in the present Experiment 2, may also have complicated interpretation of the Rock and Fleck data. For example, Rock and Fleck find significantly more *punished* responses to their A-B figure. If the subject with strongest left-pointing positional preferences were eliminated, the number of recognitions of rewarded and punished profiles would be approximately equal.

## D. DISCUSSION

In addition to supporting the main hypothesis of Schafer and Murphy, that rewards and punishments can influence figure-ground organization in an ambiguous situation, the present experiments have demonstrated that by appropriate changes in method, one can either obtain the Schafer-Murphy results or the apparently contradictory results of Rock and Fleck. The evidence suggests that the method of presentation, including such details as the type of tachistoscope used, the phenomenal size of the figures, and the degree to which the field is defined, may be important variables affecting the Rock and Fleck results. We do not suggest that there could not have been others. Perception is subtle; it may be influenced by a host of unrecognized variables. The Rock and Fleck subjects may not be comparable to the present subjects, as the former were younger and were not college students. Also, the Rock and Fleck use of symbols rather than names for the faces may have made the task more difficult. It has been suggested by Murphy and Hochberg (10, p. 340) that the motivation afforded by rewards of four cents has probably changed both qualitatively and quantitatively since 1943. Perhaps the Rock and Fleck subjects were not sufficiently motivated. Perhaps there were subtle differences in instructions that may have contributed to the variation. A discussion of each of these considerations is not warranted because the evidence is lacking. It is sufficient to note that the Schafer-Murphy results can be duplicated.

Apparently the task imposed by the projector tachistoscope method of presentation in the Rock and Fleck experiment and the present Experiment 2 makes recognition difficult. Not only does the light contrast between the interval just preceding exposure and the one-third sec. exposure of the projected image impose difficulties, but the greatly increased phenomenal size makes clear fixation and recognition arduous in the new situation. The perceptual situation of the Schafer-Murphy experiment and the present Experiment 1 with its small, circumscribed, constantly lighted field, in which the stimulus is introduced, is very different from the projector tachistoscope situation in which the field is large, and defined only for the brief interval of exposure. The shutter of the projector imposes other problems. Since it must mechanically open and close, the exposure time can only approximate the conditions of an exposure of a discrete time interval. The shutter speed is difficult to calibrate and may be subject to inaccuracy. The size of a stimulus object is a variable which affects a variety of perceptual tasks. For example, thresholds for real movement have been found to be lower for small fields and for small stimulus objects (5, p. 879) and it has been demonstrated (17) that larger ambiguous figures reverse more slowly than smaller ones. Hence, the present indications of higher recognition thresholds for the larger ambiguous faces are not surprising.

Another possible consideration suggested by Hochberg[6] is that when the size of a contour is large and the exposure time short, the possibility of its actually forming a single edge is small, and the subject may actually see a series of segments, one portion of the contour belonging to one face, the adjacent contour to the alternate face, the next, to the first face, etc. With a decrease in length of contour, this effect would probably be minimized. The longer contour could thus increase the difficulty of recognizing previously rewarded and punished profiles. On the other hand, the constantly lighted fields and the smaller figures used in the Schafer-Murphy experiment and the present Experiment 1 permitted more rapid and more "accurate" organizations.

---

[6] Personal communication.

It should be noted that the experiment by Smith and Hochberg (14) yielded significant results ($p < .02$) using a projector type tachistoscope with images 15 inches in diameter projected 84 to 96 inches from the subjects.[7] Their use of electric shock punishment and solid figures are important modifications of the procedure. For example, the solid figures probably permit an "easier" and more stable figure-ground organization. These dissimilarities prevent an adequate comparison with the present experiments.

Rock and Fleck note that the ambiguous figures in the post-training series were phenomenally different from the training faces. Reports from the present subjects support this. The authors reason that because of the guesswork nature of the task, subjects attempt to maintain consistency after an initial decision has been made. This makes statistical interpretation of the results hazardous, they state, because "differential total scores would then not necessarily imply so and so many recognitions of one face versus another, but rather consistent application of decisions as to which previously learned face most resembles which new global presentation" (12, p. 775). This argument neglects the following considerations: (*a*) the nature of the post-training task in the Schafer and Murphy experiment and in the present Experiment 1 did not require nearly as much guesswork as in the Rock and Fleck experiment; (*b*) the present experiments indicate that guesswork leads to a decrease, not an increase, in effects of rewards and punishments (note the significantly greater accuracy in recognition of test figures by those who admitted concern over the money as compared with those who did not. Also, in Experiment 1, where there were fewer mistakes than in Experiment 2, recognition of rewarded profiles was significantly greater); and (*c*) the distinction quoted above between "recognition" and "consistent application of decisions" is unclear. Perceptual differentiation and recognition may well involve a consistent application of decisions to new situations. However, the post-training series does represent a new perceptual learning situation, one in which new differentiations and stabilizations must be made, but which always bears a considerable relationship to what has gone before; in this case, the training series.

There is a great deal not yet clearly understood about the phenomenon, which only further research can clarify. We know that the "need" to receive money may be a factor in developing sets or expectancies to perceive a previously rewarded face and not to perceive a previously punished face. But there is much to be learned about the processes underlying reward and punishment. As Harlow (6) has suggested, this type of problem is far too complicated to be considered in terms of such a patently simple formulation as "drive reduction." Before a systematic theory can be developed it will be necessary to define the types of situations most conducive to the movement of cognitive processes in the direction of need satisfaction, as well as those in which information is sought on a reality basis in the organism's coming to terms with his environment. Bruner and Postman (3) have legitimately questioned the notion that perception is always wishful. Although the perceptual structure is often modified by processes conceptualized as autistic, some situations may well induce motivation to perceive realistically.

---

[7] These data were obtained from Dr. Hochberg in a personal communication.

The results from the present experiments indicate a large field of investigation in the area of individual differences in susceptibility to rewards and punishments. Not only is it necessary to determine the actual motivation of the subject, rather than assuming it, but the variation in the effects of these different situations on different subjects must be carefully studied and related to varying need constellations. It is possible that the striving for success, although very common in our culture, may have its counterpart in the individual who really does not want to win, the "masochist."

The experiments by Schafer and Murphy and by Rock and Fleck had too few subjects (five and six, respectively) to sample the individual variation, much less establish or refute conclusively the general hypothesis. Although repeated observations on a few subjects may yield a large enough number for statistical purposes, the variability between *individuals* must be considered, as well. The present experiments, although containing a larger number of subjects, may yet be found to be subject to this criticism as well.

It cannot be concluded from present evidence that rewards when not paired with punishments can influence figure-ground perception. The higher correlations between recognitions of the rewarded aspects of the two pairs and the fact that subjects were told in advance that they would finish the experiment with at least their original 90 cents, thus minimizing the influence of punishment, suggests that rewards were effective, but further research must establish this.

## E. SUMMARY

To evaluate the hypothesis, first tested by Schafer and Murphy, that autism can serve as a determinant in visual figure-ground perception and to clarify certain questions raised by a recent repetition by Rock and Fleck which did not duplicate the Schafer and Murphy results, two experiments were conducted. Subjects were presented with half-moon figures, each capable of being seen as a face, some of which were rewarded with small amounts of money, others punished. In Experiment 1 with a Dodge tachistoscope, when in the post-training series previously rewarded and punished faces were paired to share the same contour line and form reversible patterns, the subjects perceived the rewarded faces significantly more frequently as figure, thus supporting the Schafer-Murphy hypothesis. In Experiment 2, when presentations were made from a projector tachistoscope, many of the same difficulties reported by Rock and Fleck were encountered. Thus, two experimental groups were subjected to two different methods of presentation, the first resembling that used in the Schafer and Murphy experiment and the second similar to the one used by Rock and Fleck. The results indicate that the discrepancies between the Schafer and Murphy experiment and the Rock and Fleck repetition are in part attributable to the different tachistoscope introduced in the latter.

Subjects who expressed feelings about winning and losing the money in a post-experimental interview recognized significantly more rewarded profiles and had significantly fewer mistakes in recognition of test figures than those who denied concern over the money.

## REFERENCES

1. BLEULER, E. Autistic thinking. In *Organization and Pathology of Thought* (D. Rapaport, *Ed.*). New York: Columbia Univ. Press, 1951. (Pp. 399-437.)
2. BRUNER, J. S., & GOODMAN, C. C. Value and need as organizing factors in perception. *J. Abn. & Soc. Psychol.*, 1947, **42**, 33-44.
3. BRUNER, J. S., & POSTMAN, L. Perception, cognition, and behavior. *J. Personal.*, 1949, **18**, 14-31.
4. CHEIN, I., LANE, R., MURPHY, G., PROSHANSKY, H., & SCHAFER, R. Need as a determinant of perception: a reply to Pastore. *J. of Psychol.*, 1951, **31**, 129-136.
5. GRAHAM, C. H. Visual perception. In *Handbook of Experimental Psychology* (S. S. Stevens, *Ed.*). New York: Wiley, 1951. (Pp. 868-920.)
6. HARLOW, H. F. Mice, monkeys, men, and motives. *Psychol. Rev.*, 1953, **60**, 23-31.
7. LEEPER, R. W. A study of a neglected portion of the field of learning—the development of sensory organization. *J. Genet. Psychol.*, 1935, **46**, 41-75.
8. MISHKIN, M., & FORGAYS, D. G. Word recognition as a function of retinal locus. *J. Exper. Psychol.*, 1952, **43**, 43-48.

9. Murphy, G. Personality. New York: Harper, 1947.

10. Murphy, G., & Hochberg, J. Perceptual development: some tentative hypotheses. *Psychol. Rev.*, 1951, **58**, 332-347.

11. Pastore, N. Need as a determinant in perception. *J. of Psychol.*, 1949, **28**, 457-475.

12. Rock, I., & Fleck, F. S. A re-examination of the effect of monetary reward and punishment in figure-ground perception. *J. Exper. Psychol.*, 1950, **40**, 766-776.

13. Schafer, R., & Murphy, G. The rôle of autism in a visual figure-ground relationship. *J. Exper. Psychol.*, 1943, **32**, 335-343.

14. Smith, D. E., & Hochberg, J. E. The "autistic" effect of punishment on figure-ground perception. *Amer. Psychol.*, 1952, **7**, 243-244. (Abstract.)

15. Tolman, E. C. Gestalt and sign-gestalt. *Psychol. Rev.*, 1933, **40**, 391-411.

16. Wallach, H. Some considerations concerning the relation between perception and cognition. *J. Personal.*, 1949, **18**, 6-13.

17. Washburn, M. G., Maley, H., & Maylor, A. The influence of the size of an outline cube on the fluctuation of its perspective. *Amer. J. Psychol.*, 1931, **43**, 484-489.

18. Werner, H. Comparative Psychology of Mental Development. (Rev. Ed.) Chicago: Follett, 1948.

19. Woodworth, R. S. Re-enforcement of perception. *Amer. J. Psychol.*, 1947, **60**, 119-124.

*Veterans Administration Hospital*
*St. Cloud, Minnesota*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 359-365.

# RECALL OF SUCCESS AND FAILURE AS A FUNCTION OF SUBJECTS' THREAT INTERPRETATIONS*[1]

*School of Education, Stanford University*

JAMES L. GILMORE[2]

## A. INTRODUCTION

Studies of the relationship between subjects' personal interpretations of an experience and their recall of successful and unsuccessful aspects of that experience abound in disagreement. Alper (2), Glixman (7), and Rosenzweig (13) have published reviews of this literature. The conflicting evidence in this area is so extensive that two reviewers, Cameron (4), and Sears (15), have expressed doubt concerning the value of this type of research. In the present investigation an attempt was made to provide experimental controls for several factors which may account for the variations in the results of previous studies.

The first of these controls concerns the length of time *S*s spend on successful and failed tasks. Rosenzweig (11) and Sanford (14) used jigsaw puzzles as tasks in their experiments. Their *S*s were allowed to complete successfully some of these puzzles and were interrupted (failed) on others. With this procedure it would seem impossible to allow the same amount of time for the completed and the incompleted tasks. This time discrepancy would have been an uncontrolled factor in these experiments. The present study was planned so that each *S* spent equal amounts of time on successful and failed tasks.

A second control concerns the number of successful and failed tasks performed by each *S*. Glixman (8) and Atkinson (3) do not adequately control this factor. In their studies a clever, highly motivated *S* might very well have completed more than half of the experimental tasks while a dull, poorly motivated *S* might have completed fewer than half of them. This situation introduces the possibility that selected *S*s are recalling different ratios of successful and failed tasks. In the present study an equal number of successful and failed tasks were performed by each *S*.

---

*Received in the Editorial Office on June 19, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is based on a part of a dissertation submitted in 1953 to the Committee on Graduate Study of Stanford University.

[2] Now with the Department of Psychology, San Jose State College.

A third control concerns the establishment and evaluation of the degree to which each *S* perceives a personal threat in the experimental situation. Glixman (8), Rosenzweig (11), and Atkinson (3) attempt to create different levels of personal threat by giving different groups of subjects different instructions. Alper has criticized this type of procedure. She holds that some experimenters ". . . assumed that subjects adopt the particular mental set which the experimenter intends by his instructions. Yet it is by no means certain that a particular experimental *Aufgabe* can establish dynamically equivalent mental sets in different individuals regardless of what the experimenter intends" (1, p. 403). In the present study the degree of stress in the test setting was determined by each individual *S,* not by the experimenter.

## B. Problem

This study was designed to test two hypotheses.

1. The first is that *S*s who judge the test situation to be relatively nonthreatening will tend to recall more failed than successful tasks. According to Lewinian theory, tensions in the psychological field which are created by the failed tasks are not resolved in the case of the *S*s who perceive little threat. Actually, for these *S*s, these tasks are perceived as incomplete rather than failed. These persisting tensions result in a recall preference for the failed tasks. This type of mechanism has been referred to as the *Zeigarnick effect* (16), or what Rosenzweig (12) has called need-persistive reactions to frustration.

2. The second hypothesis is that *S*s who judge the situation to be relatively threatening will tend to recall more successes than failures. In Lewinian terms, the failed tasks are isolated from the psychological field and thereby are less available for recall. When the situation is perceived as threatening, memory for unpleasant, failed tasks is repressed. Rosenzweig (12) labeled this mechanism ego-defensive memory.

## C. Procedure

### 1. *Tasks*

Twenty pencil-and-paper tasks were used in the experiment. Several of the tasks were revisions of tasks used by Glixman (7). A 105-second time limit was used for each task. Ten of the 20 tasks were designed to be completed within this time limit. The other 10 tasks were designed to appear easy to complete, but actually these tasks have no solution.[3]

---

[3] The writer's dissertation (5), which includes reproductions of the 20 tasks and the complete results of this study, has been deposited with University Microfilms.

The 20 tasks were mimeographed, one to a page, and assembled into booklets with blank cover pages. The order of the tasks was random for each individual booklet.

### 2. *Administration of the Tasks*

The tasks were presented during either the first or second meeting of a general psychology class. At the beginning of the period the task booklets were passed out, and the following instructions were given:

> Please write your name in the upper left-hand corner of the cover of this booklet.
>
> These booklets contain some tasks for you to do. There is one task on each page of the booklet, and there is enough room on each page in which to put your answer.
>
> Each of the tasks will be scored on an all-or-none basis. Each task will be scored either as all right or all wrong; there will be no partial credit. In order to get credit for a task, you will have to finish all of it and get it all right. There will be a penalty for guessing, so do not guess. If you do not know an answer do not put anything down. Work as accurately as you can. Speed and accuracy have equal weight in the scoring.
>
> When the signal is given, turn the page and read the instructions at the top of the page; then begin working. Do not turn ahead to the next page until you are told to do so. Do not turn back to a previous page. Stop working as soon as you are told to stop. Stop immediately. If you finish before I say "stop" check your work.
>
> Work quickly. Do the best you can. Do not guess.
>
> Turn the page, and start working on the first task.

After 105 seconds the *S*s were told to, "Stop, and begin work on the next page." This procedure was continued for each of the 20 tasks. After the administration of the last task the booklets were collected.

A second booklet was then distributed. It consisted of a blank cover page and a mimeographed task interpretation page. The *S*s were requested to write their names in the upper left-hand corner of the cover page and then were instructed to:

> Write down the names of the tasks you remember. If you cannot remember the name of a task, write a brief description or give a short sample of it. Make your descriptions or samples brief, but include enough so that I can recognize the tasks clearly. Write the tests in the order which you remember them. Do not try to reproduce the order in which they were given.

---

Order Publication No. 5796 from University Microfilms, 313 N. First Street, Ann Arbor, Michigan, remitting $1.24 for 35 mm. microfilm or $9.90 for 6- by 8-inch photocopies.

Five minutes were allowed for this activity. Then the *S*s were instructed to turn to the second page of the booklet. On the second page was mimeographed the task interpretation text which was read aloud to the *S*s.

> One of the three following paragraphs describes the reason you were given these tasks.
>
> 1. These tasks which you have just finished will eventually be used in an experiment which is being planned. However, before the experiment can be performed it is necessary to find out how long it takes to complete each task since timing will be an important part of the experiment. You people have given us an idea of how long we should expect people to work on these tasks.
>
> 2. These tasks which you have just finished are a new "Intellectual Alertness Inventory" which we hope will predict scholastic success in college. Right now we are interested in developing norms for the test: we want to see how well college students in general can do. You people are part of a large group which will give us our standards. In the future, after the test is standardized, other students will be compared against your work. Your score will become part of your records here.
>
> 3. These tasks which you have just finished are the "Intellectual Alertness Inventory." It has been standardized, and has proved to be an exceptionally good predictive test of college grade-point average. The test was given to you because we wanted to see how well you as individuals could do. Your performance on this test will be checked against your grade-point average. We will use your score as an indication of your probable success in your future college work.
>
> *One* of these three paragraphs describes our purpose in giving you these tasks. Please indicate in the following space which paragraph you think most accurately describes our purpose. It is your opinion which we want.

Actually all the *S*s experienced similar task administration procedures, but individual *S*s gave different interpretations to the situation. According to his interpretation choice here, each *S* was classified into one of the three stress categories. Those individuals who chose paragraph one were considered to have interpreted the testing situation as a relatively neutral one. The *S*s who chose paragraph two were considered to have given some stress to the situation, and those who chose paragraph three were considered to have interpreted the situation as still more stressful. These three degrees of interpretation will hereafter be referred to as Situations *L, M,* and *H* respectively.

After the interpretation choice had been indicated the second booklets were collected, and the *S*s were requested not to discuss the tasks outside of class.

During a later class meeting a true explanation of the experiment was given. At this time only two of the 444 *S*s reported that they had suspected that some of the tasks could not be completed.

### 3. *Subjects*

The *S*s were college students enrolled in general psychology courses during the Winter Term, 1953, at San Jose State College. Of the 444 *S*s, 129 interpreted the experimental situation as Situation *L*, 171 interpreted it as Situation *M*, and 144 interpreted it as Situation *H*.

## D. Results and Discussion

For the total group of *Ss* there is a significant ($P < .01$) recall preference for successful tasks. This preference is most marked ($P < .01$) for those *S*s who indicated the stressful (Situation *H*) interpretation. It is less marked ($P < .05$) for the *S*s who gave medium (Situation *M*) interpretation, and the tendency was not statistically significant for those *S*s who gave low stress (Situation *L*) interpretations. These data are presented in Table 1. Critical ratios were computed by the mean difference technique (9, pp. 71-75).

TABLE 1

Mean Recalls of Failed and Successful Tasks, and Critical Ratios of Mean Differences

| Interpretation | *N* | Mean recall<br>Failed tasks | <br>Successful tasks | *CR* |
|---|---|---|---|---|
| Situation L | 129 | 4.48 | 4.78 | 1.46 |
| Situation M | 171 | 4.63 | 5.01 | 2.18* |
| Situation H | 144 | 4.16 | 4.99 | 4.08** |
| All Interpretations | 444 | 4.43 | 4.94 | 2.68** |

*Significant at the 5% level.
**Significant at the 1% level.

The first hypothesis, which predicted that *S*s who perceived a minimal degree of threat would recall more failed tasks, is not supported by the results. Within the structuring of the present experiment no group of *S*s exhibited the need-persistive functions which would have resulted in the preferred recall of failed tasks. Pachauri found that, "recall of incompleted tasks is greatly impaired with a lapse of time between performance and recall" (10, p. 457). It may be that the experimental structure must be such that *S*s perceive a possibility of doing further work on the failed task before these tasks are favored in recall. Atkinson's (3, p. 389) findings are suggestive of such a possibility.

The second hypothesis, which predicted that *S*s who judged the test situation to be relatively threatening would recall more successfully completed than failed tasks, is supported by the results. With increasing degrees of threat interpretation, *S*s demonstrated an increasing recall preference for the successfully completed tasks. This recall preference for the successful tasks gains increasing degrees of statistical significance with assending degrees of threat interpretations.

These results suggest the operation of the mechanism of repression. The greater the degree of personal threat that an individual *S* perceived in the experimental situation, the more likely he was to forget his failures and remember his successes.

## E. SUMMARY

A test instrument, consisting of 10 successful and 10 failed tasks, was administered to college students enrolled in general psychology courses. Procedures were followed which yielded data on *S*s recalls of successfully completed and failed tasks, and their individual judgments of the affective nature of the experimental situation.

The results indicate a recall preference for the successfully completed tasks which becomes increasingly significant with increasing degrees of threat interpretation. The higher the stated degree of threat perception which an *S* gives the testing situation, the more likely he is to forget his failures and remember his successes. This tendency suggests the operation of the mechanism of repression.

## REFERENCES

1. ALPER, T. G. Memory for completed and incompleted tasks as a function of personality: correlations between experimental and personality data. *J. Pers.*, 1948, **16**, 104-116.
2. ———. The interrupted task method in studies of selective recall: a re-evaluation of some recent experiments. *Psychol. Rev.*, 1952, **59**, 71-88.
3. ATKINSON, J. W. The achievement motive and recall of interrupted and completed tasks. *J. Exp. Psychol.*, 1953, **46**, 381-490.
4. CAMERON, N. A. Abnormalities of behavior. In C. P. Stone (*Ed.*), *Annual Review of Psychology, 1951*. Stanford, Calif.: Annual Review, 1950. (Pp. 189-206.)
5. GILMORE, J. L. Memory of success and failure in a learning situation. Unpublished doctor's dissertation, Stanford Univ., 1953.
6. GLIXMAN, A. F. An experimental study of some of the interrelationships between recall and resumption of complete and incomplete activities in neutral and stress situations. Unpublished doctor's dissertation, Univ. California at Berkeley, 1947.
7. ———. An analysis of the use of the interruption technique in experimental studies of repression. *Psychol. Bull.*, 1948, **45**, 491-506.

8. ———. Recall of completed and incompleted activities under varying degrees of stress. *J. Exp. Psychol.*, 1949, **39**, 281-295.

9. McNEMAR, Q. Psychological Statistics. New York: Wiley, 1949.

10. PACHAURI, A. R. A study of Gestalt problems in completed and interrupted tasks: Part II. *Brit. J. Psychol.*, 1935, **25**, 447-457.

11. ROSENZWEIG, S. The recall of finished and unfinished tasks as affected by the purpose with which they were performed. *Psychol. Bull.*, 1933, **30**, 698.

12. ———. Need-persistive and ego-defensive reactions to frustration as demonstrated by an experiment on repression. *Psychol. Rev.*, 1941, **48**, 347-349.

13. ———. The investigation of repression as an instance of experimental idiodynamics. *Psychol. Rev.*, 1952, **59**, 339-345.

14. SANFORD, R. N. Age as a factor in the recall of interrupted tasks. *Psychol. Rev.*, 1946, **53**, 234-240.

15. SEARS, R. R. Personality. In C. P. Stone (*Ed.*), *Annual Review of Psychology, 1951.* Stanford, Calif.: Annual Reviews, 1950. (Pp. 105-118.)

16. ZEIGARNIK, B. Über das Behalten von erledigten und unerledigten Handlungen. *Psychol. Forsch.*, 1927, **9**, 1-86.

*Route 2, Box 269A*
*Campbell, California*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 367-382.

# THE EVALUATION OF TRAINING AND TRANSFER PROGRAMS IN TERMS OF EFFICIENCY MEASURES*[1]

*Department of Psychology, Stanford University*

DOUGLAS H. LAWRENCE

## A. INTRODUCTION

Many of the decisions made in designing training programs, including those involving a transfer paradigm, are based on the principle of minimizing expenditure. The design of the program is aimed at producing the desired level of proficiency on the criterion task with a minimum outlay of time, energy, and cost. The implications of such a principle, however, have not been explicitly formulated either as they bear on the efficacy of the decisions made or on the research needed as a basis for such decisions. Consequently these decisions are tenuous. There is no certainty that they are based on the appropriate information or that when such information is available that it is weighted in an appropriate manner. The present paper attempts to remedy part of this deficiency by means of a preliminary, simplified formulation of the training and transfer problem in terms of this principle of minimization of expenditure.

In the development of the following sections a number of simplified assumptions are made about cost functions and learning curves involved in training. These assumptions are made primarily to permit a clear exposition of the methodology involved. Any specific training situation in which this methodology is used may involve a different set of assumptions about the cost function and learning curve. As a result the conclusions reached about selection procedures, the type of transfer to be desired, and the like may differ to some extent from those stated in this paper. Consequently the conclusions of the present paper are to be considered as illustrations of the type of question raised and the answers suggested by this methodology rather than as general statements true of all types of training and selection procedures.

---

*Received in the Editorial Office on June 19, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in part by the United States Air Force under contract No. AF 33 (038)-14350, monitored by the Perceptual and Motor Skills Research Laboratory, Human Resources Research Center. Permission is granted to the United States Government for reproduction; translation, publication, use, and disposal of this article in whole or in part.

## B. EVALUATION OF TRAINING

For simplicity in the presentation of the training problem, it is assumed that the criterion task is a well defined skill where the level of proficiency at any moment can be measured in terms of some performance score such as the number of correct choices, speed of reaction, or the like. The first training problem considered is that of bringing a given individual up to a specified level of performance on this task by means of a series of practice trials, each of a constant duration. This specified level of performance is assumed to be low enough so that he can achieve it with continued practice, i.e., it is below his asymptotic performance level. It is further assumed, for the purposes of this analysis, that the learning curve of this individual can be represented by a linear function. This assumption is not too limiting as long as the criterion performance level is below the individual's asymptote. Many learning curves are approximately linear up to this point.

The objective of training, as indicated above, is to bring this individual up to the specified level of proficiency with the minimum expenditure of time, energy, and money. The extent to which this can be achieved will depend upon two sets of factors: (*a*) the characteristics of the individual's learning curve, and (*b*) the nature of the "cost function" associated with this specific task. It is this latter factor that is considered below.

### 1. *Cost Function*

Each trial requires some expenditure of supervision time, of equipment, of money, and of many other factors. The efficiency of any training procedure can be evaluated only in terms of the degree to which it minimizes the expenditure of these factors, while producing the desired level of proficiency on the criterion task. For the purposes of this illustration the idea of a "cost function" will be used as though a direct outlay of money is involved on each trial. The concept, however, is much broader. It is meant to cover all possible standards, determined independently of the training procedure itself, by means of which the training procedure is to be evaluated. In some instances this cost function will represent direct expenditure of money. In others it will involve the concept of risk, i.e., the probability of damage to equipment or personnel. In still others it will involve the expenditure of time, effort, and like factors. In the most general case, the specific criterion used will be some weighted composite of all of these. To be of practical value, it is essential that these criteria be amenable to quantification.

The mathematical relationship between cost and stage of training must be determined specifically for each situation. In general, however, it is probable that cost is inversely related to the level of proficiency achieved at any given time. If for instance, in the case of pilot training, individuals begin their training on jet planes without any previous flying experience, the risk involved of damage to the equipment or personnel, the amount of supervision required, and like factors are great during the early stages of training but decrease to some minimum as training proceeds. More specifically stated, in most situations cost tends to be some decreasing exponential function of the level of proficiency obtained. In other situations, such as the laboratory, the cost may not decrease but rather be constant from trial to trial.

### 2. *Training Cost for an Individual*

For the moment, however, it is assumed that the cost per trial of training is a constant designated by *K*. The cost of training a given individual up to a criterion performance level, where this criterion level is designated as *Pc,* can be determined if his initial performance level on this task prior to any training, designated by *Po,* is known as well as his rate of learning, designated by *B*. Then the number of training trials he requires is shown by Equation (1).

$$N = \frac{Pc - Po}{B} \qquad (1)$$

Each of these trials requires an expenditure equal to *"K"* so that the cost of training this given individual is

$$\text{Cost} = K \frac{Pc - Po}{B} \qquad (2)$$

This equation can be simplified by defining *A* as equal to 1/B, i.e., as the reciprocal of the rate of learning, and by defining *M* as equal to Po/Pc, i.e., by expressing the initial level of performance on this task as a fractional part of the criterion level of performance. With these substitutions, the cost of training this individual is

$$\text{Cost} = K\,A\,P_c\,(1 - M) \qquad (2a)$$

*a. Reducing training costs.* Even in this oversimplified formulation of the problem, several implications for the minimization of the cost of training are emphasized. Assuming that the cost per trial, *K,* is fixed, the total cost is directly proportional to *Pc* and *A*, and is inversely proportional to *M*. It is apparent that if *Pc,* the criterion level of proficiency, is set too high in any training situation, it will increase the total cost considerably when the cost per trial is at all large. This underlines the necessity of using a precise and rational method of determining the value of *Pc*.

One approach to the problem of when to discontinue training is in terms of cost functions. Consider the case in which each training trial results in some improvement in performance but at the same time involves a constant investment of time, effort, and money. On the other hand, if the individual is placed on the job at this stage of training, there are definite costs involved largely in terms of risk or probability of damage to equipment and personnel. Thus there are two cost functions to be considered, that characteristic of the training program and that characteristic of the job. Both are functions of the level of proficiency obtained at any given time. The optimum point at which to discontinue training is the point at which an additional trial of training does not produce a sufficiently large change in performance level so that the resulting reduction in the job cost function compensates for the actual cost of that training trial. Consequently, this problem of determining the point at which to discontinue training is itself a problem in minimizing expenditure, and can only be solved when specific cost functions have been determined for both the training program and the actual job for which the individual is being trained.

Once $Pc$ is fixed the cost of training an individual can be reduced by changing either the value of "Po" or the value of $A$. Changes in the value of $Po$, initial performance level, normally will have relatively little influence on the cost of training. It is true that the cost of training varies directly with changes in the value of $M$, the ratio of $Po$ to $Pc$. As $Pc$ will tend to be large compared to $Po$ in most situations, however, even a substantial change in the absolute value of $Po$ will have a relatively small effect on $M$. Consequently in order to reduce training costs, the procedures used must be directed primarily at changing the learning rate rather than the initial performance level.

*b. Decreasing cost function.* Equation 2a was developed on the assumption that the cost per trial was constant regardless of the level of performance. As indicated previously, a more realistic assumption is that the cost per trial varies inversely as the level of performance, and probably approaches some asymptotic value. As a first approximation to this state of affairs, it is assumed that the cost per trial is linearly related to the mean performance level during that trial and that the slope of this function is negative. This relationship is shown in Equation 3 where "Ci" is the cost of the *ith* trial, $Pi$ is the mean performance level during that trial, $K$ is a constant characteristic of the situation, and $k$ is the slope coefficient.

$$\mathrm{Ci} = \mathrm{K} - \mathrm{k\,Pi} \tag{3}$$

If an individual is trained from a starting performance of $Po$ up to the criterion level of *Pc,* it can be shown that the average cost per trial during this training is represented by

$$\overline{C_1} = K - \frac{k}{2} P_c(1 + M) \tag{4}$$

Multiplying this equation by Equation 1, the number of training trials required, gives the total cost of training this individual.

$$C_{\text{total}} = AP_c(1 - M)\left[K - \frac{k}{2}(1 + M)\right] \tag{5}$$

In this type situation, changes in the value of *M* play a more important rôle than in the previously described situation. A given percentage change in "M" produces an even larger percentage change in the total cost than does the same percentage change in *A*. The reason for this is that a decreasing cost function has been assumed so that an increase in the initial performance level, *Po,* eliminates some of the most expensive trials during training. This effect is greatest when both the slope of the cost function, *k,* and the criterion performance level, *Pc,* are large. In evaluating this effect, it must be borne in mind that it is the value of *M* and not the change in "Po" that is involved. Consequently, even in this type situation it would appear to be more feasible to reduce the cost by selecting for the rate of learning than to select for *Po.*

*c. Concept of error.* In the discussion of both of the above situations, one important variable has been ignored. The true objective of training when specifically stated is to permit, *with a known probability,* the statement that a given individual has reached or exceeded the assigned performance level of *Pc.* This, however, cannot be judged on the basis of the individual's performance on any given trial. His performance scores are subject to "error" or variability. The "true" level of performance at any moment can be judged only by taking the entire learning curve into account, i.e., it must allow for the variability of the performance scores about the regression line, the variability of the regression coefficient, and the variability of the intercept value (in this case *Po*). As all of these sources of variability are some function of the variability of the performance scores about the regression line, it will be assumed as a first approximation that their collective effects are proportional to this latter source of variability, hereafter designated as *E.* Consequently, the individual must be trained to an obtained performance level of "Pc" plus *dE* in order to justify a probability statement that he has reached or exceeded *Pc.* The value of the multiplier, *d,* in the above expression is determined by the level of probability decided upon. In the following examples it is given the value of *1*, which implies that the chances are about 85 in 100 that this individual has reached or exceeded *Pc* at the end of training.

The concept of *E* merits some discussion. This measure is usually thought of as representing error of measurement in the sense of being due to the unreliability of the measuring instrument. When such is the case, this value can be determined for the specific task under consideration, and then is constant for all individuals. Equally well, however, a part of it may be due to the characteristic behavior of a given subject, i.e., it might well be a measure of the subject's own unique variability or unpredictability from trial to trial. It must then be estimated for each individual. For this reason *E* will be retained as an identifiable part of the formulation.

As used here, *E* is conceived of as primarily increasing the amount of training necessary because of the difficulty of determining the true value of *Po* and of determining when the individual has actually reached *Pc*. Its influence on *A* is disregarded on the assumption that the amount of training involved is sufficiently great to give a reliable estimate of the rate of learning. Consequently, if *L* is defined as the ratio, *E/Pc*, and incorporated into Equation 5, the expression for the total cost of training becomes

$$C_{\text{total}} = AP_c(1 + L - M)^2 \left[ K - \frac{k}{2}(1 + L + M) \right] \qquad (5a)$$

A change in the value of *L* in this equation has an effect analogous to that of a change in *M*, as already described, but its influence is somewhat less in that it is not necessarily concentrated at the beginning of training where the cost per trial is greatest. It should also be noted that as *L* approaches *M* in magnitude, changes in *M* have a decreasing influence on the cost of training. The implication is that individuals must be selected in terms of *L* as well as in terms of *M* and *A*, or if the former is not possible, selection in terms of *M* will have relatively little influence. Again it should be remembered that the discussion is in terms of *L*, the ratio of *E* to *Pc*, and not in terms of the absolute magnitude of *E*.

### 3. *Average Training Cost for a Group*

So far only the factors influencing the cost of training a single individual have been considered. In a practical situation, however, all members of a large group must be trained to the criterion level of performance. As this group must be considered a random sample from some population, the individuals will tend to differ widely in terms of their initial performance levels, their rates of learning, and the size of their error terms. The problem then becomes one of minimizing the average cost of training per individual. An expression approximating this mean value is given in Equation 6. The barred letters represent average values for the population and the *V*'s are coefficients of variation.

$$\begin{aligned}\overline{C}\text{ per individual} = {} & K\bar{A}P_c[1 + \overline{L}(1 + r_{AE}V_AV_E) - \overline{M}(1 + r_{AP_0}V_AV_{P_0})] \\ & - k/2\bar{A}P^2{}_c[1 + 2\overline{L}(1 + r_{AE}V_AV_E) + \overline{L}^2(1 + V^2{}_E + 2r_{AE}V_AV_E) \\ & - \overline{M}^2(1 + V^2{}_{P_0} + 2r_{AP_0}V_AV_{P_0})] \qquad (6)\end{aligned}$$

Any conclusions drawn from this specific formula are limited, of course, by the specific assumptions underlying the cost functions and learning curves of this illustration. Nonetheless, this expression points up some aspects of the problem of training that normally are overlooked. For instance, the average cost per subject is determined not only by the mean values involved but also by the variances and intercorrelations of initial performance level, rate of learning and error. Whether or not in any specific instance these intercorrelations are sufficiently great to merit consideration must be empirically determined. If they are of considerable magnitude, then training and selection procedures must take them into account. Even when they are negligible, however, the variance measures of individual differences cannot be ignored. The concept of individual differences, taken in conjunction with the idea of a cost function, makes it possible to formulate more explicitly the type of consideration that should govern decisions made in designing training programs.

### 4. *Training Methods and Selection*

There are essentially three major ways in which training costs can be reduced: by improving training techniques, by selecting the individuals to be trained, and by utilizing transfer of training procedures. As transfer is discussed in the second half of this paper, only the implications for training and selection will be considered at present. Most research on improving training techniques, such as studies of the influence of spaced versus massed practice, of motivational level, of knowledge of results, and the like, are primarily concerned with modifying the rate of learning. The objective of selection procedures, however, is less clearly defined as a rule.

It has been assumed, so far, that each individual in the group is trained until he reaches the criterion level of performance. This, of course, requires different amounts of practice for each individual because of differences in initial performance level, rate of learning, and size of error term. In most large scale training programs, however, all individuals are given the same amount of practice irrespective of individual differences. The implicit assumption is that a standardized training program is less expensive than one that must be varied to allow for such differences. But if all individuals are brought up to the required level of proficiency in a standardized program, the majority of them are being overtrained. This adds to the total expense of the program. Consequently, to justify this overtraining, it is necessary to show that the average cost per trial is considerably less under the standardized program than under the individually tailored one. This differential in average cost between the two methods of training is most likely to occur for rapidly decreasing cost functions, for then the overtraining will be at a relatively low cost per trial. It is less likely to occur for relatively flat cost functions and for very heterogeneous groups.

If an individually tailored program is adopted, consideration of individual differences suggests other ways of reducing the cost by modifying the training procedures. As indicated previously, studies on training methods concerned with spaced versus massed practice, motivation level and the like are primarily directed at modifying the average rate of learning. Normally, however, a standardized procedure for the spacing of practice and like factors is adopted for all individuals of the group. It is doubtful if this produces the optimum effect for each individual. The optimum value for one individual might well differ considerably from the optimum value for other individuals in the group depending upon the variability in levels of performance at any given time, the rates of learning, and other characteristics of the individuals involved. Consequently, the maximum possible change in rate of learning is not being achieved. If research were to show that individual differences in reactions to these training procedures were of considerable magnitude and could be predicted prior to training, allowance for such individual differences in the training program adopted might permit a worthwhile reduction in the average cost.

Another type of training situation frequently encountered is one in which selection procedures are involved along with a standardized training program. Instead of attempting to train all individuals in the group to the specified level of performance, only those who meet the criterion after a specified number of trials are accepted while the remainder are rejected. The basic problem is to select individuals prior to training in such a way as to maximize the number who will meet or exceed the criterion at the end of training.

When the selection device is a test validated against a pass-fail criterion on the specific training task under consideration, it can be shown that its effectiveness is largely due to the fact that it is measuring the rate of learning. In most situations a pass-fail criterion is a rough index of the final performance level of an individual after a fixed amount of practice. This final performance level can be expressed as follows where *Po* is the initial performance level, *N* is the number of training trials, and *B* is the rate of learning.

$$P_f = Po + NB \tag{7}$$

If this final performance level is correlated with a test, *Y*, the resulting correlation can be expressed as follows.

$$r_{P_fY} = \frac{Nr_{BY}\sigma_B + r_{P_0Y}\sigma_{P_0}}{\sqrt{N^2\sigma^2_B + \sigma^2_{P_0} + 2r_{BP_0}\sigma_B\sigma_{P_0}}} \tag{8}$$

When the two variances in the above equation are roughly of the same magnitude and *N* is large, practically the entire correlation is due to the relationship between the rate of learning and the test scores. Thus selection tends to be largely in terms of rate of learning. Possibly the efficiency of this type of procedure could be increased if the test were designed specifically to measure this aspect of performance.

The above formulation is strictly true, however, only if the criterion performance level, *Pc,* is below the asymptotic performance level for any individual in the group. If *Pc* is set so high that a large percentage of the individuals in the group do not achieve it even with continued practice, then tests validated against an acceptance-rejection criterion are selecting both in terms of rate and in terms of asymptotic performance level without distinguishing beween these two sources of failure. Each of them, however, has a distinguishable influence on the cost of training. For instance, if a slow rate of learning is the basis for rejection, the sizeable investment that has already been made in partially training the individual can be salvaged in part with further practice, but if the rejection results from a low asymptotic performance level, none of the training investment can be salvaged. In situations where a limited number of candidates are available but a fixed number of trained individuals are needed, the ideal type of selection from the point of view of cost would be one in which all individuals with low asymptotic values could be rejected prior to training and the needed number of individuals selected from those of the remainder that had the highest learning rates. This procedure demands two types of selection devices, one to predict the rate of learning and a second to predict the asymptotic level of performance. This is on the assumption that these two measures are not highly correlated in the population.

The above example indicates a general approach to the problem of training in terms of selection for homogeneity of subjects. As indicated previously, the argument in favor of standardized training programs is that the cost per trial tends to be less than under the individually tailored program. At the same time, it is wasteful in that it overtrains some individuals and rejects some of the partially trained ones. If, however, groups of individuals are selected who tend to be relatively homogeneous with respect to initial performance level, rate of learning, size of error terms, and performance asymptote, there is a minimum of overtraining or loss through rejection. This is on the assumption that the amount of training given is specified in terms of the average values selected. In addition, there is the savings due to a standardized training program. Once the mean values for the population on this task are determined, it becomes possible to specify the degree of homogeneity within a group and the number of different training programs needed in order to produce a given number of trained individuals at a minimum cost. All this is programmatic, however, until selection instruments are specifically designed to separate out the various aspects of performance, primarily the rate of learning and the asymptote levels, and secondarily the initial performance levels and error terms.

### C. Evaluation of Transfer

In addition to selection devices, one of the most important ways that the cost of training can be minimized is by employing transfer of training situations. The assumption underlying the use of the transfer paradigm is that an individual can be brought to the desired level of performance on the criterion task more easily, or with less cost, if he first is trained on a preliminary task, than he can be without such preliminary training. When the cost of the preliminary training is taken into account, however, the assumption becomes somewhat more tenuous, for it then must read that with transfer the cost of preliminary training plus the cost of criterion training is less than the cost of criterion training alone without transfer. Consequently, the problem of this section of the paper is to determine the circumstances under which this assumption is justified and the type of information that is necessary o justify it.

It is apparent from Equation 6 that prior training on a preliminary task will reduce the cost of training on the criterion task only if it changes the average value of either $A$, $M$, or $L$ or some combination of these. This gives rise to two general problems: first, to determine which of the above aspects of criterion task performance it is most desirable and feasible to modify by transfer in any specific situation; and, secondly, to determine the nature and amount of the preliminary training that will give the optimum amount of transfer.

The first of these problems has been the basis for most of the theoretical interest in transfer experiments. A considerable amount is known about the relationships between two tasks, such as stimulus and response similarity, necessary to produce transfer in terms of changes in the initial performance level, i.e., the value of $M$. Only casual attention, however, has been given to the relationships necessary to produce transfer in terms of changes in the rate of learning. As will be demonstrated later, transfer of this latter type is probably the single most important way of reducing the cost of training through preliminary practice.

### 1. *Amount of Preliminary Training*

Whenever there is a transfer effect, the amount of this effect must be some function of the number of preliminary training trials. The basis for this statement is the following. In order to change the values of $A$, $M$, or $L$ on the criterion task, the preliminary training must have changed the behavior of the individual in some manner. Consequently, some preliminary training is necessary before any such change occurs, and the extent of this change can be expected to vary with continued training. At some stage of preliminary training, however, little if any further change in behavior, and, therefore, in transfer, will occur; in fact, the initial changes even may be cancelled out by overtraining effects.

If this description tends to hold true in the general case, then the function relating changes in the criterion task parameters to the amount of preliminary training will be curvilinear. Consequently, there should be some specific amount of preliminary training that will minimize the total training cost. If individuals are not trained sufficiently on the preliminary task, there will be a loss due to the necessary additional training on the criterion task. If they are overtrained on the preliminary task, the additional change in the criterion parameters will not be sufficient to compensate for the cost of this overtraining. The basic problem, therefore, is to determine the relationship between changes in $A$, $M$, and $L$ on the criterion task and the number of practice trials on the preliminary task, $n$, in such a way that a value can be assigned to $n$ that will minimize the total cost of the preliminary plus criterion task training.

The problem when specifically formulated takes the following form. The cost of training a single individual on the criterion task, when the cost per trial is assumed to be constant, is shown in Equation 2a. This holds when no preliminary training has been given. This cost is designated by $C$. If instead the individual is trained on the criterion task after preliminary training so that transfer occurs, the cost is somewhat less and can be designated by $C'$. The savings, i.e., the difference between $C$ and $C'$, is equal to $dC$. This savings is brought about because the preliminary training changes the values of $A$ and $M$ in Equation 2a to the values $A'$ and $M'$. Defining $dA$ as equal to $A - A'$ and $dM$ as equal to $M' - M$, so that positive values of each represent positive transfer, the expression for the change in cost of training as the result of transfer is as follows.

$$dC = K\,Pc[dA(1-M) + (A-dA)dM] - K'n \qquad (9)$$

The second term on the right, $K'n$ is the cost of the preliminary training necessary to produce the transfer. It is assumed that the individual is given $n$ preliminary trials each of which costs an amount $K'$.

### 2. *Types of Transfer*

This equation points up two extreme types of transfer that can occur; transfer in terms of changes in the rate of learning as represented by $dA$ and transfer entirely in terms of changes in the initial performance level as represented by $dM$. When transfer is of the former type involving only a change in the rate of learning, the reduction in cost on the criterion task is directly proportional to the change in $dA$. When transfer is of the latter type, however, involving only a change in the initial performance level, the reduction in cost is not proportional to the change in $dM$. It depends equally upon the value of $A$, the reciprocal of the rate of learning. If the rate of learning is slow, so that the reciprocal is equal to or greater than unity, then a change in $dM$ can be effective in reducing the cost. Otherwise such a change has a proportionately smaller effect on cost than does a comparable change in $dA$. The possibility of using transfer involving only changes in the initial performance level to reduce cost is further limited when it is remembered that "M" is defined as equal to $Po/Pc$. This means that "dM" in the above equation is equal to $\frac{1}{Pc} \times dPo$. When $Pc$ is a large number, as it will be in most training situations, a unit change in $Po$ will produce only a fractional change in "dM" and consequently in the reduction of cost. This means that even large changes in the initial performance level of the individual as the result of transfer may have only negligible effects in reducing the cost of training. Thus a transfer situation designed to produce a change in the rate of learning is more desirable than one that is designed to produce a change in the initial performance level. Unfortunately most research has been based implicitly on the opposite assumption.

The final term, $K'n$, in the above expression must always be kept in mind as it represents the expenditure on the preliminary training. To be efficient as a training method, transfer situations not only must reduce the cost of training on the criterion task, but this reduction must be sufficiently large to compensate for the cost of preliminary training. This is most likely to be the case whenever the cost per trial on the preliminary task is relatively small compared to the cost per trial on the criterion task, or when the optimum transfer effect can be produced in relatively few trials. If both $K'$ and $n$ become large, it is doubtful if a transfer paradigm can be justified.

Two aspects of the usual training situation have been ignored in the above equation. First of all no allowance has been made for "E," the error term on the criterion task, nor has the inverse relationship between cost and performance level been taken into account. It is apparent in Equation 7a that a decrease in the size of $E$ as a result of preliminary training would reduce the cost of training on the criterion task. While no attention has been given to this possibility in terms of research, it is potentially as important a source of savings as changes in $Po$. Consequently, the change in "L" brought about through transfer, here designated as $dL$, is defined as the difference, $L - L'$, where $L$ is the ratio of the error term to criterion performance level prior to preliminary training and $L'$ is the same value after such training. At the same time a cost function is assumed that decreases linearly with the obtained level of performance on the criterion task. The resulting expression for the savings in the cost of training a given individual is

$$\begin{aligned} dC &= KPc[dA(1 + L - M) + (A - dA)(dL + dM)] \\ &- k/2\ Pc^2[dA(1 + 2L + L^2 - M^2) + 2(A - dA)(dL + LdL \\ &- MdM) - (A - dA)(dL^2 + dM^2)] - K'n \qquad (10) \end{aligned}$$

Again it is apparent that changes in $dA$ have a direct effect on reducing the cost whereas the effect of changes in $dL$ and $dM$ are inversely related to the rate of learning. With a decreasing cost function, however, the influence of $dM$ is slightly enhanced in that it tends to eliminate the early and, therefore, most expensive training trials.

In actual practice it is not this reduction in cost for a single individual that needs to be maximized in order to obtain the most efficient transfer situation. Rather it is the average of such expressions for all members of the group being trained that must be maximized. It is apparent that this average reduction in cost will be a complex expression determined by three different sets of factors: (*a*) the average values of $dA, dL$, and $dM$, (*b*) the average values of $A$, $L$, and $M$, and (*c*) the intercorrelations and variances of these terms. In addition, there will be an expression for the cost of the preliminary training. Assuming that the intercorrelations and variances have a negligible influence and that the average values of $A$, $L$, and $M$ have been determined for the population by means of control studies, the problem of maximizing this average cost reduction is one of finding the optimum values for $dA$, $dL$, and $dM$.

As indicated previously these latter always can be expressed as some function of $n$, the number of trials of training on the preliminary task. These expressions will tend to be of a non-linear form such that beyond some point in preliminary training little if any change in $dA$, $dL$, and $dM$ will occur. At the same time the cost of the preliminary training will continue to increase. Consequently, if these expressions in terms of $n$ are substituted into the formula for the average reduction in cost, it should be possible to find some value of $n$ that will maximize this average.

### 3. *Influence of Individual Differences*

The situation is complicated by individual differences. There is no a priori reason why a given amount of training on the preliminary task should have the same transfer effect for all individuals. Transfer occurs only when certain types of changes take place in the individual's behavior during the preliminary training. While the extent of such behavior changes will be some function of $n$, they will vary also with certain individual performance characteristics during preliminary training such as the rate of learning, $b$, the individual's initial performance level on the preliminary task, $po$, and his variability from trial to trial, $e$. These individual differences must be taken into account when specifying the relationships of $dA$, $dL$, and $dM$ with $n$.

The measurement of these individual characteristics during preliminary training poses a problem. The present state of theoretical knowledge provides little guide as to which changes in behavior during preliminary training are related to changes in the parameters of the criterion task. While certain aspects of behavior during preliminary training are observed and recorded, they may have little relationship to other concomitant changes in behavior that actually produce the transfer effect. In such instances the changes in the criterion task parameters can be related to $n$ but not to any of the other performance characteristics on the preliminary task.

It will be assumed here, however, that the observed changes in individual performance during preliminary training are related to the changes in the criterion task parameters. These relationships can be expressed in the following general form.

$$\begin{aligned} dA &= F_1\,(m, b, P_o, e) \\ dM &= F_2(m, b, P_o, e) \\ dL &= F_3(m, b, P_o, e) \end{aligned} \qquad (11)$$

Until these functions are empirically determined, the most that can be postulated about them is: (*a*) each is zero when *n* is zero, (*b*) each has some maximum value or at least an asymptote as *n* increases, and (*c*) the value of "n" producing this maximum varies with the values assigned to *b, po,* and *e,* the performance characteristics on the preliminary task.

The fact that the value of *n* which maximizes the reduction in cost is a function of the preliminary performance characteristics implies that the practice of giving all individuals in a transfer situation the same amount of preliminary practice is inefficient. This is especially true when the individual differences are at all marked. Some individuals are undertrained with the result that little if any transfer takes place. Others are overtrained so that the cost of their training goes up without any corresponding increase in transfer. The alternative is to adjust the number of preliminary training trials for each individual in terms of his performance characteristics on this task. This demands knowing the *b, po,* and *e* for each individual. It would seem feasible to make such estimates in a sequential manner as the individual proceeds with the task.

An additional benefit in terms of selection might result from these sequential estimates during preliminary training. It was pointed out in the discussion of training procedures that the use of homogeneous groups on the criterion task permits all the benefits of a standardized training program and at the same time prevents waste due to overtraining or to the rejection of partially trained individuals. If the values of $A'$, $L'$, and $M'$, i.e., the rate of learning, the error, and the initial performance level of the individual on the criterion task after preliminary training, can be predicted from a knowledge of his *b, po, e,* and *n* on the preliminary task, then it becomes possible not only to assign the value of *n* to each individual that produces the maximum reduction in cost but at the same time select out homogeneous groups for training on the criterion task. This selection advantage would fully compensate for the added expense of giving individual attention to each subject during preliminary training.

When the problem of transfer is viewed in the general form outlined in this paper, it is possible to list some rough criteria to guide the search for the most efficient types of preliminary training when the nature of the criterion task is specified. First of all, the cost per trial on the preliminary task must be relatively low compared to the cost per trial on the criterion task, or else maximal transfer effects must be produced in relatively few trials. Secondly, of the various forms of transfer that are possible, it would appear to be most advantageous as a general rule to emphasize transfer that increases the rate of learning on the criterion task. This is the most effective way of producing considerable reductions in the overall cost of training. Thirdly, the preliminary training should be of such a nature that the performance characteristics of individuals during this training relate to the amount and type of transfer effect that occurs. Otherwise considerable efficiency in transfer is lost because of the overtraining or undertraining of individuals on the preliminary task. Finally, it would be desirable to have these same performance characteristics relate to the absolute values of the criterion task parameters as they exist after preliminary training. This would permit the use of the preliminary task as a selection device as well as a training situation.

*Department of Psychology*
*Stanford University*
*Stanford, California*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 383-387.

# RESEARCH WITH NEUROPSYCHIATRIC SAMPLES*

*Department of Psychology, Michigan State College*

GERALD F. KING[1]

The inadequacy of the present neuropsychiatric diagnostic categories is practically common knowledge. In psychological research, a major trend has been the tremendous increase in studies utilizing neuropsychiatric subjects; and such research, if it is to be effective, must cope with the present diagnostic classifications. The mishaps of research involving neuropsychiatric subjects have been pointed out by Kelly (9), Windle (14), and others. It is felt that this is such a crucial research problem that further discussion is warranted.

The importance of the problem is illustrated in the literature by numerous studies which reveal a discrepancy between the knowledge about diagnostic categories and the experimental methodology employed. Such studies frequently utilize diagnostic status, at least in part, as the "independent variable." For example, predictions are made about schizophrenic behavior, with other nosological groups or "normals" serving as controls or standards of comparison. The discrepancy is revealed when the subjects are indiscriminately selected for the experimental sample, as long as they bear the diagnosis "schizophrenia." The major concern seems to be to equate the experimental group with comparison-groups on such variables as age, education, intelligence, and so forth. Schizophrenia, itself, is hardly questioned. The practical outcome of such research is aptly described by Thorne, as an editor, who questions the value of "manuscripts which speak vaguely about experimental samples consisting of 100 psychoneurotics or 25 psychopaths since such terms are so semantically indefinite as to have only slight differentiating value" (13, p. 89).

For the sake of simplicity, the inadequacy of the neuropsychiatric classifications can be viewed in terms of *reliability* and *homogeneity*. The poor agreement among clinicians in assigning the mentally ill to different diagnostic groups has long been noted in practice, and this unreliability of the diagnostic categories has been confirmed by empirical studies (1, 2, 3, 11). The rea-

---

*Received in the Editorial Office on June 19, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to thank Professor Albert I. Rabin for his interest and suggestions.

son seems to be that the criteria are vague, and, in addition, cut across diagnostic lines. In the matter of composition, diagnostic categories are broad and heterogeneous, rather than restricted and homogeneous. Such a diversity of behavior is subsumed under diagnostic groups that one gets the impression, for example, that the difference between two schizophrenics can be as significant as the difference between a normal and a schizophrenic. The uncritical acceptance of these diagnostic categories in experimentation implicitly assumes that they are clear-cut, homogeneous entities similar to typical medical diseases. The results seem to be reflected in the small group differences usually reported and the all too frequent failures of cross-validation studies.

Although both reliability and homogeneity are important factors, homogeneity seems to be more closely related to the structure of psychodiagnosis. It is recognized that reliability can be influenced by different conditions of homogeneity. However, with a more rigorous definition of diagnostic criteria, reliability would undoubtedly be improved, but homogeneity would not necessarily be appreciably affected. Making the diagnostic categories more homogeneous would entail revising the entire taxonomic structure. For the research worker, the heterogeneity embraced by the various diagnostic labels seems to pose the more difficult problem.

The influence of heterogeneous conditions has been illustrated in several recent studies. Kantor, Wallner, and Winder (8) divided their schizophrenic subjects into "reactive" and "process" groups on the basis of case history data. It was found that the Rorschach protocols of the two groups were significantly different in terms of being judged "psychotic" or "nonpsychotic." In another study, Windle and Hamwi (15) investigated the relationship between test performance and prognosis in two groups of psychotics similar in diagnosis but different in length of illness. Performance on a "complex reaction time" test was positively related to prognosis in the early group, while this relationship was reversed for the chronic group. Using various criteria to establish groups of "short term" and "long term" schizophrenics, Rabin, King, and Ehrmann (12) found no significant differences between a normal group and the short term schizophrenics on measures of vocabulary efficiency, comprehension, and communication. The obtained differences were between the latter two groups and the long term schizophrenics. In summary, these findings indicate that within diagnostic categories, especially schizophrenia, what might be called "chronicity" is an important variable for predictive purposes.

Some hypothetical data, which parallel some actual findings, can be used to show how heterogeneity is related to research results in a more subtle manner. The subjects consisted of 20 normals and 20 paranoid schizophrenics, the two groups being equated on the usual variables. The dependent variable can be called "false perceptions," a measure derived from the tachistoscope. The hypothesis was that the paranoid schizophrenics would produce more false perceptions than the normals. The performances of the paranoid schizophrenics and normals revealed means of 14.1 and 10.6 and sigmas of 5.7 and 2.6, respectively. Statistical tests indicated that the schizophrenics showed both significantly more false perceptions and more variability than the normals. However, when five paranoid schizophrenics who showed marked "confusion" were eliminated, the mean and sigma of this group dropped to 10.7 and 2.4. As can be seen, this wiped out any differences between the groups. Thus, the previous significant differences were due to "confusion" or at best "confused paranoid schizophrenia."

Greater performance variability for neuropsychiatric samples does not appear to be an infrequent research finding. Perhaps, in many cases it could be accounted for in a like manner. This example would also indicate that even the more restricted diagnostic subtypes (in this case, paranoid schizophrenia) may present problems of homogeneity.

A possible rationale for accepting diagnostic categories and utilizing heterogeneous groups could be that general diagnostic factors might be discovered in this way. The above example would certainly argue against this approach. It is interesting to note in this connection that in a factor analytic study of schizophrenics, Guertin (5) was able to identify several types of schizophrenia, but no evidence was found for a general schizophrenic factor. It should also be pointed out that with such diagnostic innovations as "psuedoneurotic" schizophrenia (7), the indication is that the diagnostic category of schizophrenia is becoming even broader and more heterogeneous.

Although this discussion of research with neuropsychiatric samples has not been an extensive one, it is felt that it has certain implications in terms of suggestions for future research. The following three avenues for investigation, which are actually highly interrelated, are suggested.

1. The most obvious and, at the same time, the most ambitious type of research would be that aimed at revising the present inadequate diagnostic classifications. According to Thorne (13), this is the greatest need of the fields of clinical psychology and neuropsychiatry. This type of research could utilize various orientations, ranging from differential etiological studies to attempts at improving the present practice of describing patients according to symptoms. Currently, such investigators as Eysenck (4) and Wittenborn, Holzberg, and collaborators (see 6) are interested in this problem.

2. On a more modest level, heterogeneity, itself, seems to provide a frame of reference for fruitful study. In the past, such research has usually been bound to the present diagnostic taxonomy, as attested by the usual practice of comparing diagnostic subtypes (e.g., hebrephrenic vs. paranoid schizophrenia). It is suggested that other than the standard criteria of heterogeneity should be searched for and utilized in experimentation. It seems relevant to note that the distinction between process and reactive schizophrenia has been in the literature, either explicitly or implicitly, a good many years without receiving much attention from psychologists. Other neglected psychopathological differentiations are undoubtedly available. Findings in this area could contribute evidence that would be utilized in modifying the structure of the diagnostic system.

3. As previously inferred, much of the present research is broad in scope in the sense that the focus is upon general psychopathological categories (e.g., schizophrenia, psychoneurosis). These investigations could usually be methodologically improved. The experimental subjects should be described more rigorously, preferably in operational terms with specific variables or functions. The samples should also be more homogeneous. This procedure would tend to make the selection of subjects independent from the present diagnostic classifications and would provide for more clarity in the interpretation of the results. As pointed out by Windle: "If investigations were carried out under homogeneous conditions we could eventually build up a composite picture, an accomplishment much less likely to result from studies in which these variables are uncontrolled" (14, p. 473). In short, as much care should be given to the selection of the subjects as that given to the formulation of hypotheses, construction of experimental techniques, and other aspects of research.

A current trend that is especially compatible with the preceding viewpoint can be seen in the research approaches which are beginning to treat diagnosis as a multi-dimensional concept (e.g., Wittenborn). The quantification of symptomatic behavior is a significant step for experimentation, especially when it is considered that psychopathological symptoms extend across diagnostic lines. It should also be pointed out that personality, in general, is multidimensional. It would seem that better predictions could be made, as suggested by Lorr (10), if control were exercised not only in regard to psychopathological dimensions but also in regard to other relevant personality variables. The influence of such a non-pathological variable as motivation, as expressed by the subject's coöperation, is offered as a simple example.

## SUMMARY

The inadequacy of the present diagnostic categories was discussed in terms of its relationship to research utilizing neuropsychiatric samples. Diagnostic categories were described as unreliable and heterogeneous. The heterogeneous behavior subsumed under the various diagnostic labels was emphasized as a factor contributing to inconclusive research results. Implications in terms of suggestions for future research were given.

## REFERENCES

1. ASH, P. The reliability of psychiatric diagnoses. *J. Abn. & Soc. Psychol.*, 1949, **44**, 272-277.
2. DOERING, C. R. Reliability of observations of psychiatric and related characteristics. *Amer. J. Orthopsychiat.*, 1934, **4**, 249-257.
3. ELKINS, F. Specialists interpret the case of Harry Holzer. *J. Abn. & Soc. Psychol.*, 1947, **42**, 99-111.
4. EYSENCK, H. J. The Scientific Study of Personality. New York: Macmillan, 1952.
5. GUERTIN, W. H. An inverted factor-analytic study of schizophrenics. *J. Consult. Psychol.*, 1952, **16**, 371-375.
6. HOLZBERG, J. D., & WITTENBORN, J. R. The quantified multiple diagnostic procedure in psychiatric classification. *J. Clin. Psychol.*, 1953, **9**, 145-148.
7. HOCH, P., & POLATIN, P. Pseudoneurotic forms of schizophrenia. *Psychiat. Quart.*, 1949, **23**, 248-276.
8. KANTOR, R. E., WALLNER, J. M., & WINDER, C. L. Process and reactive schizophrenia. *J. Consult. Psychol.* 1953, **17**, 157-162.
9. KELLY, E. L. Theory and techniques of assessment. *Ann. Rev. Psychol.*, 1954, **5**, 281-310.
10. LORR, M. The classification problem in psychopathology. *J. Clin. Psychol.*, 1953, **9**, 143-144.
11. MEHLMAN, B. The reliability of psychiatric diagnoses. *J. Abn. & Soc. Psychol.*, 1952, **47**, 577-578.
12. RABIN, A. I., KING, G. F., & EHRMANN, J. C. Vocabulary performance and levels of communication in schizophrenia. (To be published.)
13. THORNE, F. C. Editorial opinion. *J. Clin. Psychol.*, 1953, **9**, 89-91.
14. WINDLE, C. Psychological tests in psychopathological prognosis. *Psychol. Bull.*, 1952, **49**, 451-482.
15. WINDLE, C., & HAMWI, V. An exploratory study of the prognostic value of the complex reaction time tests in early and chronic psychotics. *J. Clin. Psychol.*, 1953, **9**, 156-161.

*Psychology Department*
*VA Hospital*
*Battle Creek, Michigan*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 389-419.

# THE MULTIPLICITY OF THE NORMAL EGO STRUCTURE*

*Portal House of the Chicago Committee on Alcoholism*

STANLEY ROSENMAN[1]

## A. INTRODUCTION

An unquestioned postulate of present-day personality theory, given credence by the psychoanalytic and Gestalt schools alike, is the unity of structure of the healthy personality. Unity of personality structure is here defined as a tightly organized hierarchy in which all peripheral parts are pale reflections, instrumental extensions, or, at best, subordinate satellites of the single dominant theme. Distinctiveness is theoretically permitted a given section of the total organization only as this state represents an actualization of some minor aspect of the ruling theme. In the context of the currently predominant phenomenological approach, unity is chiefly defined as the essential similarity, the uniquely characteristic nature of all the individual's self-experience. This definition represents a retreat by the centralists from the former peripheralistic belief in behavioral unity as defined by an outside rater. It was at the latter point that unity theorists, admittedly erroneously, had earlier argued the issue.

It is the purpose of this paper to demonstrate that unity may be a structure that accommodates mental illness, and that multiplicity of structure may accompany, and indeed be an intrinsic aspect of, the healthy individual. Here we shall attempt to give support to the hypothesis of the manifoldness of the normal individual's self-experiences and the multitudinousness of the derived stable self-representations. It is proposed to distinguish between the adaptive and maladaptive forms of both singularity and multiplicity of the personality, and more particularly, of the ego structure. The exact types of multiple personality structures that are correlative to pathology will be elaborated. Ways in which conflict between different structures of the personality may be avoided or resolved will be discussed. Finally, the implications for personality structure of an ideology of multiplicity as opposed to an ideology of unity will be taken up.

---

*Received in the Editorial Office on June 20, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The writer wishes to acknowledge the helpful suggestions of Dr. E. Hanfmann, Dr. S. Star, and Dr. M. Zimmerman.

## B. HISTORY OF THE UNITY-MULTIPLICITY PROBLEM

The ego as a construct was developed in both academic personality theory and in psychoanalysis to serve analogous purposes. Once the Gestaltists (24, 25, 26) had argued for the hierarchical organization of the personality, they had to designate explicitly that structure to which was attributed the domination of the hierarchy by its effectively guiding and organizing all the symbolic processes of the organism. In psychoanalytic theory (10, 11, 8) the ego similarly serves as the synthesizer of the conflicting demands made upon the individual, and particularly, to transmute id processes—non-rational, seething, ignorant of dubiety or negatives or logical imperatives—so that both in their expression on a conscious level and in their expression in behavior they were orderly, rational, and, above all, "integrated." It is perhaps ironic that one questions whether the ego, a construct introduced as the agent of the supposed universal unity and integration of the normal individual, may not be better conceptualized as a multiple structure.

A prominent and still unresolved problem in ego theory today is the relation of the ego as subject, i.e., as agent of psychological processes, to the ego as object, i.e., as the reflexive product of the ego's activity. The second problem with which we are largely concerned is still largely embedded in ego theory by virtue of the general consensus in the field. It is the problem of unity versus multiplicity as an attribute of the ego and as a resultant of its functioning. We have raised the first problem, i.e., ego as subject and ego as object, as it is upon the postulates we offer in its solution that we shall develop our approach to the second question.

Some theorists, e.g., the orthodox psychoanalysts, deal almost exclusively with the ego as the subject or agent of psychological processes, relegating the ego as object, i.e., as self-image or self-concept, to a residual status. Other theorists, e.g., James (19) and Rogers (31), make the ego as object the cornerstone of their theory, ignoring the rôle of ego as subject. Theorists who deal with both aspects of the ego have varying and usually unclear treatments regarding the interrelations and extent of overlap of the ego as subject and as object (e.g., 4, 27).[2]

Despite the explicit consensus upon a unified ego as an effective agent of integration of the psychological processes of the individual, there are many hints that this is not quite the case. For example, in psychoanalytic theory, there are constant insinuations of the multiplicity of the normal personality and of the normal ego. The division of the personality into the three mental institutions, id, ego, and superego may be translated into three different types of self-other experiences widely held by different individuals. The supposed reifications of these institutions and their interactions further indicate their personified nature. Within the superego itself there are the ego ideal, the private conscience, and in times of social crisis, the group conscience. There is a splitting of the ego due to the ego's different object identifications that is not considered pathological by psychoanalysis. Analysts (34) further argue that in therapy there is a normal dissociation of the ego as the therapist allies himself with the reality-oriented aspect to do combat upon the repressive, distrustful, and affective aspect of the ego. The symptoms derived from different libidinal levels in every neurotic and the potentiality of the use of regression as a defense in the healthy individual are further indicative of the multiplicity of the ego structure insinuated by Freudian theory (2). Federn (6), who constantly reiterated that the normal ego is a cohesive unit, still stressed its being differentially cathected and variably active under different conditions.

---

[2] Federn (6), too, does not appear to deal explicitly with this problem. He simply lists the different frameworks he uses without explicitly indicating their interrelationships.

Close scrutiny of other personality theorists, e.g., Angyal (5), would demonstrate that despite their explicit embrace of the ideology of the unity of the normal individual, there is implicit allowance for multiplicity of structure. In contrast to the above insinuations of multiplicity, James' clear depiction of the many selves of the normal individual has been regarded as simply another aspect of the quaint idiosyncrasy of this charming writer. It is of further background interest to us that sociologists emphasize the many social rôles of the individual. For the person, as he plays out the rôles prerequisite to the continued functioning of the society, takes positions at a number of points in the social structure. Should one assume that at the core of every social rôle is a concept of the self together with the pertinent significant other's legitimate and perhaps illegitimate expectations of one's behavior in that rôle, the implication here too is of a multiplicity of the self-structure.[3]

## C. BASIC POSTULATES

Before presenting our position on the unity-multiplicity issue let us briefly state our prerequisite assumptions on the relation of the ego as subject to the ego as object.[4]

---

[3] For a fuller presentation of the historical background and discussion of these two difficulties see Sections *B* and *F* of (32).

[4] In the following, the words "representations," "schemata," "definitions," "structures," "images," "pictures," and "percepts" as referring to the self and the significant other are used as roughly equivalent. They represent the individual's interi-

The basic principles guiding and regulating the behavior of the individual in his life situations are his definitions of himself as these are modified by the definitions of his significant others and the perceived interrelationships of self with others. The individual acts in the social scene in a manner designed, if only unconsciously, to give realization to his self-other themes. Should an individual appear to himself to possess a given trait, he will act so as to provide for himself data that will reinforce this self-representation. Significant others will be chosen on the basis of the "fit" of their qualities with those of one's other-concepts. Actions directed towards them will further provoke their conformity to the existent other-images.

Perhaps the chief test of these assumptions is whether the individual gives expression in action to negatively-toned themes. It is a much more readily acceptable presumption that one will desire to actualize positively-toned themes. Yet the individual may act out negative themes for a number of reasons:

1. One may seek to work out or master a negative theme. Still if the negative aspects of the theme are strong enough when given expression, the behavior simply provides further reinforcement to the negative theme (e.g., a combat nightmare).

2. One may act out a negative theme as a way of avoiding becoming consciously aware of the theme, e.g., acting out defiance against authority so as not to have to face consciously one's own need to castrate authorities. But even if one does not become conscious of it, the data provided by acting out the negative theme must reinforce it.

3. Themes, although negatively toned, still provide the individual with a basic frame of reference which he will fight to maintain. They are a conceptual home base.

4. As will be seen later, a negative theme may represent a split from a potentially integrated theme, providing the necessary balance to very positive themes. The individual "accepts" the negative theme as the penalty for the positive theme. One will choose, for example, to suffer inferiority feelings so that one may continue to enjoy cherished narcissistic fantasies.

5. The individual may give expression to his negative self-image to see if others still continue to love him despite his acting his "worst self." He may test the other's love.

6. The individual may unconsciously identify with a sado-masochistic

---

orized and relatively stable organizations of his appraisals of his experience. They are the meanings signified by the individual of himself, his environment, and their interrelationship.

significant other whom one consciously hated, yet unconsciously loved. By suffering in the same manner in which he did, one may expiate one's guilt for hating the other, and one may express one's love. It is suggested, for example, that this is one factor determining the alcoholic adjustment of many individuals who had an alcoholic parent.

We have thus argued that the constellation of self-other themes determines the manner of active functioning attributed by other theorists to the highly unified ego in its rôle as subject or agent. These themes are also the self-reflexive product of this manifold functioning, i.e., they are both ego as object and ego as subject.

It is in this context that we shall attempt to demonstrate that the normal individual may possess a number of self-other structures, i.e., a multiple ego.

A repeated social situation results in experiences approximately unique to that situation. For example, the actual achievement level, the achievement level relative to what is expected by oneself and by others, the actual achievement relative to that of one's peers in any situation may diverge from that of other situations for the individual. The significant others of various situations tend to be different persons, perhaps variant types of persons, with dissimilar demands on and evaluations of the individual. Even if they are the same individuals, their expectations of and attitudes toward the subject may be contradictory or distinctive for the separate situations. A mother, for example, may act very differently toward her child when they are in company than when they are alone. Both the individual and his significant others tend to evaluate him in different and even in opposite ways for the nonidentical situations, e.g., "John is a good clinician but a poor theorist." The individual consequently has contrasting reactions to the different evaluations. There is a variance in the quality, including the rhythm, the intensity, and the interweaving of pleasures and frustrations from situation to situation. Thus the individual learns to perceive himself and others differently in his different situations.

Variations in the figure-ground relationships of the self-other structures may be determined in part by the given social rôle the individual is currently taking, the given significant others with whom he is presently interacting, the particular goal after which he is striving, and the dominant physiological need. There is an intrinsic rhythm, however, to the figure-ground alterations. Each self-other theme may be expected to rule the stage of the personality both qualitatively and quantitatively according to some of its structural dimensions. Thus, the self-other theme will determine and provoke social situations that will permit it to actualize itself. It may even evoke the regnancy of a particular physiological need. For example, the individual unconsciously feeling insecure and unloved may seek a sexual partner with the conscious self-experience only as that of being in sexual heat. Once in figure, the particular self-other theme leads to a differential access in consciousness of given ideals, memories, fantasies, moods, and even of capacities from the total store potentially available to the individual.

## D. NEGATIVE ASPECTS OF MULTIPLICITY OF PERSONALITY STRUCTURE

It is not difficult to see the reasons why the united personality has been held up as an ideal. Social critics of the American scene have long denounced a culture that prizes an atomistic multiplicity: disparate bits of knowledge, of possessions, of relationships. The multiple personality, conceptually assimilated to this trend, has been regarded as invariably fragmented into unorganized parts. The shattered pieces were regarded as simply resonating differentially with the demands of the individual's many social rôles. The multiple personality was an unpleasant and confused cacophony.

As the psychotherapist saw multiplicity of structure only in the context of pathology, he contributed to the argument that multiplicity is a cancer which feeds upon the valid and limited purposes of the organism. He pointed to the individuals who give mechanical responses dependent upon the stimuli fed into them, to the sorely conflicted patients who have many opposing definitions at different levels of the same situation, and to the precious dilettantes who confuse productivity with a dabbling in all fields. The latter, the professional smatterer in relationships and fields of study, has indeed provided the unity theorists with a legitimate whipping post. For the sciolist develops to a limited degree a wide number of self-other structures. The full development of any one structure is avoided as he fears that he will not make the grade should he invest a major part of his cathexis in any one self-other structure. He may fear that he will not be amply rewarded. He may prefer to dream of accomplishments in all fields or many relationships without risking the possibility of having the cherished fantasies burst by reality. He may prefer the passive absorption rôle to that of making an active contribution in a relationship or a field, thus leaving every field or relationship after absorbing and before giving. Dilettantism remains pleasantly non-satiating. One is like a Don Juan who, by avoiding becoming involved in any relationship, need not face the basic themes of one's personality. Often the dilettante has a split self-picture. There is a narcissistically omnipotent self-percept that refuses to accept the organism's realistic limitations or the environmental obstacles to achieving a goal. This self-percept is then balanced by a self-percept of despairing inadequacy and fear of failure.

## E. POSITIVE ASPECTS OF UNITY OF PERSONALITY STRUCTURE

Certainly an ideal attainment, for which some kinds of healthy personalities strive, is the ability to consciously perceive simultaneously all the self-other structures of the personality, their interrelations with each other, one's energic system as it overflows into the psyche, and the stimuli from the outside world. Such a conscious perception may provide the basis of a rational preening and compromising system that would tend to integrate the different self-other structures into a cohesive and non-conflictful structure. Although such a structure need not be *logically* interlocking, it may be possible for such an individual to consciously discover a basic self-other theme from which all others may be logical derivations or, at least, logically related. Such an integrative perception may permit all parts to communicate with each other. It may allow the individual to rationally select the appropriate self-other structure to best meet a given situation. Every aspect of the personality would have its "proper" place in the personality; the individual would in a sense always "know where he was" within himself.

However, this ideal is not quite what many unity-theorists have in mind. They apparently claim that an individual experiences himself in the same way regardless of the situation in which he is. The individual is represented as a rather simplified unit so that it is no difficulty for those who are not pathological to experience their totality, and it is this same totality which supposedly enters every situation. It would appear to this writer that these theorists may have identified the sense of conscious "myness" that accompanies all the self-experiences of the individual upon their reaching consciousness for the entire range of distinguishable self-experiences of the individual.

Some unity-theorists who proclaim a dignity of man based upon the invariability of the individual's self-experience, no matter what his surrounding, perhaps mistakenly equate two genotypes. One is the personality structures wherein the picture of the other is more dominant than the self-picture in the determination of the individual's behavior. This state leads to personalities that differentially but mechanically respond to whatever stimuli their significant others provide them. The second genotype is a normal multiplicity with one's self-concepts in full communication with and control of one's psychic energy.

Horney (17) demonstrates this customary attitude towards unity and multiplicity when she discusses a dependent individual who does whatever the members of her present group do. Horney thinks thereby to exemplify loose integration. Yet it would equally appear that the patient is well integrated about the theme of imitative dependency. Value judgments about unity seem to have determined Horney's choice of classification in this case.

### F. Negative Aspects of Unity of Personality Structure

Healthiness of the personality structure need not mean the oppressively strict and tight binding of the personality in a single theme. The Fascist organized about the theme of sado-masochism, the oral person oriented about a sucking dependency upon others, the acquisitor trampling upon others in his single-minded race to amass possessions, are all exemplars of the warning that a rigid cohesiveness, a singular unity of life purpose, should not be equated with the ideally normal.

Unity, like multiplicity, is simply a possible framework of the personality structure. Before evaluating unity as positive, one must ask the contents of what is being unified. One must further know the process of unification: whether by throwing out via denial and repression large and vital areas from the conscious personality or through a creative perception of the interrelationships of the self-other structures. For many personalities are integrated or non-conflicted by virtue of the poverty of personality themes, as well as of their immediate environment. Here, unity may be accompanied by such traits as non-creativity, stereotypy, and conformity.

Earlier, we pointed out that anxiety may lie behind the promiscuous development of self-other themes of the dilettante. Anxiety may equally well be the motive force of overspecialization, i.e., permitting one self-other structure to take over the entire personality. One may feel comfortable or secure only with the dominance of one self-other structure. All others teem with threat.

It would appear legitimate to ask that we look behind the concepts of integration and cohesiveness when used as criteria of mental health and see to what extent anxiety has determined either integration or multiplicity.

It should further be noted, as Rank (29) has pointed out, that an individual attributes meanings to differentiation and integration. For example, wholeness of personality may be feared as it represents death or primitive dependence on the mother; multiplicity may be feared as it represents departing from the mother to the differentiation of adulthood with death as the inevitable conclusion. However, multiplicity should perhaps not be opposed to wholeness, but rather to singularity. Wholeness may best be contrasted with partialness, and even mutilation.

One meaning unity theorists have given to integration has been integrity. Correspondingly, multiplicity has been equated with duplicity and irresponsibility. Yet the very opposite may be the case. An ideology of integration may serve as the rationalization for surrendering the manifold development that life demands. Desperate seekers after unity as an ideal have often attributed such unity to an organization or principle to which humiliating subservience is thereafter given. An ideology lauding unity may undermine the possibilities for the permissive looseness of personality which tends to seek an accepting pluralism in the social structure and a democratic federalism in the political structure.

It is unfortunate that some psychologists have confused their lay public by their emphasis upon integration and consistency. Their assumptions may easily be misunderstood as a demand for uniformity and sameness. Psychologists here represent to the individual his significant other authority figures who demand a rigid homogeneity from situation to situation. This exact reduplication of behavioral characteristics in one's different situations permits the authority figure a more accurate pigeon-holing of the individual, a greater predictability of his behavior, and a readier holding of him as responsible for any contrariety.

We note here with some hesitancy that those psychologists who have held aloft the banner of integration and of unity, of the importance of the whole over the part, have come mainly from a nation whose dedication to these same principles has had disastrous results. Indeed, these very psychologists have been among the first to suffer, as well as to do vigorous battle against the totalitarian national structure. It is ironic, therefore, that they espouse aspects of the same ideology as an ideal for the personality structure.

An additional factor in the customary overemphasis upon unity of self may be the result of having only one body. The body is, however, elaborately differentiated into given systems which are differentially active at various times. Further, we may not properly generalize from the body to the self-structure. Although one brings the same body to the different situations—and even this is questionable in terms of different tensions, postures, expressions, patterns of functioning, etc.,—one still has a multiplicity of experiences in different situations. One organizes one's psychological selves about these repeated distinctive experiences. The mind is characterized by its emergent values over bodily dynamics, and its multiplicity of structures may be one of these.

## G. POSITIVE ASPECTS OF MULTIPLICITY OF PERSONALITY STRUCTURE

The multiplicity of self-other structures may, even when leading to conflict, provide the individual with the prerequisites for a fuller awareness of his potentialities, and if he wishes, their realization. Perhaps only by having a multiplicity of self-other structures may the individual become aware of his individuality and be released from the sense of inevitability in permanently acting out any given self-other schemata. Each self-other theme is like a small case of valuable culture contact, permitting the individual to be more aware of all the self-other assumptions upon which he operates, and potentially to revise them in a direction he consciously chooses.[5]

A multiplicity of self-other themes permits gradual change of the personality structure. The organism can slowly "feel out" a self-other structure, and invest it with large amounts of energy only after it is found satisfying. There need not be an all or none change in which a structure, kept primitive and unconscious because of one's ideology of unity, takes over the entire personality in a "conversion" experience—usually to the detriment of the individual and the society. The multiply-structured individual may use the existent personality structure as a base from which to launch new and better ways of relating to the world. Perhaps the pathologically dissociated personality has the greatest chance of arising, not in the rationally multiple individual, but in one embarked upon the crusade to fulfill the ideology of single self-hood.

If one had a completely unified personality, its invariable expression in action would tend to rapidly lead to oversatiation with consequent anxiety and impoverishment. Oversatiation would also ensue should all one's social situations require a relatively strict unity of behavior. Should the social structure refuse one the privilege of expressing one's single self-other theme, the tightly united individual's equilibrium would be quickly undermined.[6] Presumably, this collapse would not occur in the case of an individual with differentiated self-other structures who could then turn with vigor to another social field of action for expression, creativity, and perhaps as a base upon which to launch a drive for the renewal of conditions that permitted the expression of the first outlawed personality structure.

One may note that even in apparent conflict of personality systems, only the expression of one may fully free the other for expression, e.g., "Love is never so sweet as after a quarrel." When one self-other theme represents a reaction-formation against the other, its expression may be prerequisite to the second structure's liberation. Often, only after an expression of love may one become aware of one's dislike of the other, or may one rebel against the other.

---

[5] A therapy situation would be impossible unless the therapist has different self-other assumptions than does the patient. This difference alone enables the therapist to question the patient's assumptions.

[6] Thus the depression with its loss of jobs caused the detached migrant worker to lose his one connection to the larger society. He was quickly demoralized and rapidly sunk to an apathetic Skid Row existence. See (38).

Theorists regard a divided individual as a sick one. Yet singularity of purpose may result in a rigidity that defeats itself. In combat, one's forces are often divided so as to better accomplish a mission. Only rarely, however, do the two forces end up by shooting at each other. In like manner, an individual may engage in more "integrative" behavior, i.e., organizing one's actions so as to most efficiently reach a desired goal, by having a multiplicity of self-other structures upon which to draw to meet the different obstacles before the goal. Perhaps the academician has confused the normal individual's external behavior in a laboratory short-range means-end situation with his internal structure. One may be "divided" and because of it, a more competent human being.

Ideally then, every self-other structure in the healthy individual is characterized not so much by its ability for logical interlocking with the assumptions underlying all other self-other structures, but rather by its feeling that (*a*) the only limit for its expression is a largely self-imposed one: a system may be given expression for just so long before being satiated and/or (*b*) it is limited no more than any other structure in the interests of some overall tendencies of the organism, e.g., security, whose achievement subserves all the structures of the personality and/or (*c*) its own limitations in serving the pleasure needs of the organism in contrast to the potentialities for pleasure of of other systems qualify the extent to which it comes into figure. Each self-other structure "feels" that nothing in the personality is basically antagonistic to its view or satisfaction. It neither threatens nor is threatened by other self-other structures in the personality, and at the same time, no fear lies behind its not communicating with any other self-other structure.

### H. VARYING ORGANIZATIONS OF SELF-OTHER STRUCTURES IN DIFFERENT INDIVIDUALS

The personality structures of different individuals may best be conceptualized as having different organizations of self-other themes. Some may have one dominant theme to which all others are subordinate. Here the peripheral or surface self-other structures may play only very slight variations on the central theme of the personality. Other individuals may have two equally dominant central systems which alternate in coming into figure. Some personality structures may be conceived as having several self-other structures, of roughly equal dominance. Other individuals may have two or three dominant and extensive self-other themes. Other peripheral themes may then be subordinate to one or more of these dominant themes. At the same time there may be a number of other dominant but much less extensive themes. Still other individuals may have a large number of little related or organized self-other structures. Many schizophrenics tend to be characterized by an instability that prevents the firm organization of any self-other structure.

Unfortunately, the assumption by psychological workers that multiplicity appears only in pathologically dissociated personalities has led to very little attention being paid to the different kinds of organization of basic themes in different individuals. The determinants of distinctive kinds of personality structure and their significance for different patterns of functioning remains an area open to empirical investigation.

To exemplify a personality organization that appears to have two dominant self-images, we may quote from Redding in *On Being a Negro:*

> From adolescence to death there is something very personal about being a Negro in America. It is like having a second ego which is as much the conscious subject of all experiences as the natural self. It is not what the psychologists call dual personality. It is more complex and, I think, more morbid than that. In the state of which I speak, one receives two distinct reactions—the one normal and intrinsic to the natural self; the other, entirely different but of equal force, a prodigy created by the accumulated consciousness of Negroness.

The book reviewer continues:

> "He then cites an incident to drive home his point. As a normal human being he was moved to pity when he saw a woman, drunk or sick and clad only in a slip, stagger into a neighboring backyard in Louisville on an extremely cold day. On the other hand he felt a gloating satisfaction that she was white" (21).

Perhaps Redding labeled this state more morbid than a dual personality because of the anger- and guilt-ridden content of the self-picture organized around Negroness, and because the amount of choice he had in its creation was minimal. However, as there is communication between the two self-other systems in Redding's case, and as both systems are conscious, despite the intense disagreement between the assumptions underlying both, this should be considered objectively as much less pathological than an actual case of dual personality with its greater dissociation between the self-other structures.

Not only may the individual have a multiplicity of self-other structures, but within any given self-other theme one may normally distinguish the threads of a number of sub-themes. For example, one patient in his relationship to a female had three separable perceptions of the girl about which rotated different sub-structures:

1. The girl was perceived as a nagging mother urging him on to greater maturity. The subject here perceived himself as infantile, and mortally afraid of maturity and responsibility. He consequently hated his partner for provoking anxiety in him.

2. The girl was a seductive bed-mate whose accomplishments he admired and liked. Here he felt pride in his sexual adequacy, and in being a desired physical object.

3. The girl was an affable, mature, and intellectual companion. In this substructure he perceived himself as a promising young genius through whose eyes the girl might look upon the world with unusual insight.

Most interesting is that the patient reported three distinct physical images of the girl that corresponded with each of the above perceptions. One cannot say these were three completely different structures to which the percept of the girl was assimilated, as the patient was aware of many of the connections between the themes. Yet each theme continued to play a part in the relationship. Perhaps three different well-established self-other schemata became loosely connected by the relationship.

## I. FAULTY ORGANIZATIONS OF SELF-OTHER STRUCTURES

We have maintained then that the healthy individual may have a number of organized schemata of self-in-relation-to-other structures. It is not the multiplicity of self-other schemata which determines an individual's maladjustment but rather the manner in which the multiplicity is organized as well as the content of the different structures. We shall examine only the problem of the organization of structures here and omit study of the content. The maladjusted individual may have a number of faulty modes by which self-other structures are related to each other. There is some overlap in the different categories.

1. *Two or more self-other structures of an individual that should be carefully differentiated may remain very similar.* An individual may evaluate all others and act toward them in the same way, perceive others as acting toward and evaluating him regardless of who the other is, and evaluate himself in relation to others in the same way and with exactly the same criteria regardless of the other with whom he is interacting. One male patient had similar self-in-relation-to-other-structures vis-a-vis pretty girls and male authority figures. Only from these two groups who were perceived as aloof and forbidding did he expect rejection; and only in interaction with these two groups was he insecure. This client did differentiate other groups very finely. Another example of the failure to differentiate two structures is the

mother who plays a lover rôle toward her own child. Yet another is the employee who assimilates a realistically relaxed superior to a stern father-image.

Generalized themes of some individuals may interweave virtually all the self-other structures in the individual. One client, a college student, reported feeling rejected when his three-year-old cousin playfully told him to leave her house. Despite the obvious dismay of the youngster when the client did leave, he reported still feeling hurt. The client of course had severe problems about rejection.

It should be noted that some reality-congruent and "adequate" themes may be distinctive and yet subordinated to an over-all theme. For example, a supervisor may desire to manage all employees so as to obtain the best performance from each. He may still handle each one very differently depending on the particular person. Still this multiplicity of sub-schemata would result from the unity of the more general purpose of using the others as instruments for his goal.

Some individuals may live in a pale colorless world in which everyone, including the self, lacks all idiosyncrasy. Everyone is "nice" or "all right" or "O.K." This type of individual may equate evaluating the other with aggressing against the other. As all functions interlaced with aggression are repressed, so is any keen sizing-up of the other. Insight is the forbidden fruit. Some individuals may fear disequilibrium following a close examination of their psychological world.

2. *There may be pathological dissociations of self-in-relation-to other structures.* That is, a structure which should be united to best grasp the data upon which it is based is split into two or even more divisions, each of which selects a resonant part of the data provided by one's behavior. For example, there may be a structure wherein the self is perceived as inferior and in which the subject relates submissively to a powerful significant other. This structure is balanced and supported not by the objective values of self and other and their situation, but rather by another estranged personality structure in which the self is perceived as superior, the self relates aggressively to the perceived other and actively depreciates the qualities of the other. Freud (13) has well described the split pictures of the pure mother and the earthy whore which coerce many men into impotence with wives assimilated to the virginal mother-image.

3. *An organization that overlaps the preceding type is the building up of self-other structures in defense against other self-other structures.* The defensive structures may then occupy extensive areas of the conscious personality

without having deep emotional roots, other than the need to ward off another self-other experience. Individuals with this type of structure may continue to act out shallow selves, never becoming really acquainted with themselves or with others with whom they might build meaningful relationships. They fail to maximize their possibilities. The potentially vital self-other structures are kept from consciousness and motility. For example, an individual may maintain in a largely unconscious state the self-other structures that deal with the interaction of self and opposite-sexed and same-aged persons. This prevents the subject from working out the feared relationship so that it may give him satisfaction.

The extensive, non-intensive system often derives a major part of its energy from the self-other system that is being kept from consciousness. An individual may defend against a self-other system marked by his sadistic reaction to others by a consciously polite self-picture with correspondingly changed significant others. The surface conscious system may then use the sadistic urge to defend itself against the now buried sadistic system. However, the defensive devices available to any conscious self-other schemata are by no means limited by the need matrix of the buried system.

Many patients provide striking examples of the deep sense of alienation from one's self that occurs with extensive surface self-other structures that lack deep roots of feeling in the personality. Often both the very positive and very negative nuclei of the self-image, when the images are so split, are largely repressed and the individual remains vaguely conscious of his objective qualities in between the two extreme self-images. When one or the other repressed image breaks through to consciousness, the individual reacts with still greater confusion as to who and what he is.

4. *The self-other systems may conflict for dominance in the same situation.* In the above case of the split wife-image, the individual may vacillate between perceiving the wife as a taboo mother-figure or as a whore, and be unable to integrate the picture of the wife into a realistic unit. Thus actions toward the wife also remain divided. Generally, one or both structures are unconscious in pathogenic conflict.

5. *One self-other system may have too great centrality and extensity in relation to the subject's station in life.* For example, a forty-year-old male may be concerned only with his mother.

An individual by choice or necessity may be in close interaction with a significant other who can little satisfy his needs. For example, an individual may seek to master the trauma of a rejecting narcissistic mother by winning a narcissistic woman for his wife. It is just such a woman, however, who can least give him the warmth that he craves.

6. *The individual may confuse the sense of myness or otherness that belongs to different constellations of attributes.* A self-concept may lose the feeling of myness and be perceived as an image of an other. On the other hand, an other-concept may become imbued with the sense of myness. The latter process is also an aspect of normal growth. The evaluation of it as pathological requires the judgment that the individual is thereby divorced from his own genuine interests.

7. *The boundaries between two systems may be overly fluid.* A facetious example is the calling of one lover by the name of another. The schizophrenic personality organization is characterized by inconstancy and weakness of the self-other structures' confines. Material from one field flows into another where it does not belong.

8. *The boundaries between self-other structures may be too rigid, leading to insufficient intercommunication.* An individual may continue to treat his boss on the golf links with the same circumspections as in the office—and to the great embarrassment of the boss. An individual when in a situation with significant others of two different categories may act most ineffectively because of an inability to integrate two rigid self-other structures into a functioning whole. Rigidity of a self-other structure in the personality may lead to the assimilating of others to it when the potential objective interpersonal relations with the other need not have been confined to the narrow and frustrating limits set by the particular self-other structure. Although a particular group of the individual may welcome a wider and more spontaneous contribution from the individual, the rigidity of the boundary of self-in-relation-to-the-particular-groups may prevent the individual from drawing upon other resources of his personality when in the group.

9. *A self-other structure may have the other-percept more central than the self-percept.* Here the other-percept separates the more peripheral one of the self-percepts from the accurate description and fulfillment of his own needs. A number of clients deal complainingly with a multiply-organized personality structure. Critics may state that the inferred multiple structures do not belong to any of the faulty organizations listed above, but are nevertheless tied up with the maladjustment of the individual. Let us examine two examples.[7]

---

[7] The quotations are taken from protocols of the University of Chicago Counseling Clinic.

One young female client is trying to break away from her husband and speaks as follows of his differentially negative effect upon her:

> Of course the most miraculous thing I could hope for is that he would voluntarily leave. I don't expect that would straighten everything out. I mean . . . the energy and power and capability that I feel most of the time I'm away from him . . . well, it doesn't just go away. It must still be there, when I'm with him, but it doesn't make things possible.

And again in the next interview:

> . . . whenever I have left him then I feel strong energy and things are clearly outlined and I can see exactly what areas I want to operate in and exactly how I can try to make him see where he's wrong on certain things, and exactly what I want to do about it, and then when I'm with him again, I try to search back and try to find those things about which I felt so strongly before. And they are not clear at all. They're vague and I can't apply myself to them.

In the following interview she states: "I've just given up all reason, all power of concentration, every time I'm faced with him."

And here is another client speaking:

> And what was running through my mind was that, uh, it isn't uh just one person whom I admire but, however I do, uh well, I, I often have a feeling of going out to someone and wishing that I were like that person and it isn't always the same kind of person. And I, I just got the feeling the other evening as I sat at home by myself that I, I guess I try it out in different directions because of that towards someone whom I admire. I guess I see that more than I have in quite a while but at the same time it's something that's going on all the time, so that, uh, I'm quite a different person with different people.

The writer would maintain that it is not the multiplicity of structures that is the pathogenic agent here. Rather, at the center of the self-other structure, not the self-concept but the concept of the other controls the energy available to the structure. The other-image is then projected upon external others with the resulting other-centeredness of the behavior.[8] The last client, in addition, appears from her total protocol to have a large number of disorganized shifting structures little related to each other, and completely dependent upon the person with whom she is.

---

[8] There was recently an advertisement in a rather enlightened magazine that attempted to exploit the frequent existence in the personalities of our society of an other-concept more central than the self-concept. It read:

Are *You* an unknowing
VICTIM OF HYPNOTISM
Millions are! !

and then a statement of what to do about it. See (22). Variations on this theme are frequently found in popular mystery and adventure stories. Some aspects of the current witch hunting mania are also an expression of this theme.

She exemplifies the individual whose self-evaluations in his different rôles are determined completely by the significant others of the complementary rôle. We insist, however, that it is one thing to have the differing self-evaluations of different situations entirely dependent on the evaluations of the self by the significant others of that situation, and another to have differing self-evaluations that accord with one's differing objective qualities in different situations. The latter is, however, admittedly partially determined by the significant others of the situation.[9]

As suggested earlier, since therapists tend to deal only with neurotics, they come across the multiplicity of structures only in the context of sickness. This may mislead them into thinking that all multiplicity is therefore evidence of pathology. The academic psychologist who sees mainly normals, and these superficially, is limited in his view by the single self-other structure he aroused in the normal vis-a-vis himself and his simple laboratory situation. He then attributes a specious unity to the normal.

## J. BALANCING SELF-OTHER STRUCTURES

Dissociated self-other structures were listed as one of the modes of faulty organization and is elaborated here as it is a focal point of convergent if somewhat conflictful theories of the pathologically disunited personality. We will therefore further examine pathological self-other structures that support and balance each other, each being a split from a potentially single and objective percept of self-in-relation-to-another. It will be assumed that the given potentially whole self-other structure of a personality has split into two halves. However, it should be noted that a self-other structure may be easily fragmented into more than two parts.

Freud (9) wrote regarding the sexual instincts, "Whenever such an impulse is formed in the unconscious which can be paired with a contrast, one can regularly demonstrate that the latter, too, is effective."[10] Abraham (1) has described how narcissistic the self-depreciating melancholic really is.

---

[9] See the analogous approach though from a different framework of Riesman (30).

[10] Freud (12) in 1909 discussed the ambivalent structures of an obsessive patient. He wrote as follows:

"His (the patient's) attitude towards the lady—a compound of tenderness and hostility—came to great extent within the scope of his conscious knowledge; at most he deceived himself over the degree and strength of his negative feelings."

Freud continues:

"The conflicts of feelings in our patient which we have here enumerated separately were not independent of each other, but were bound together in pairs."

Adler and Horney (18) are two of the many writers who have discussed the growth of a glamorized self-image in reaction to feelings of lack of worth. Freud (13) first discussed a split picture of a significant other in dissociated images of the pure virginal mother and the prostitute that arise from the troubled perception of the mother's relation to the father. Jung (20), who made the balancing of opposites quite central in his theory, has written: "It is always true that too much upon one side brings about, in the unconscious, a too-much upon the other." These views are expressed most clearly by Perls (23) who argues that wherever there is a plus perception of an object there is invariably a compensating minus perception. Thus to compensate the Southerner's perception of the Negro as lazy, incompetent, etc., the Southerner has a myth of the Negro's great sexual potency.

It would seem legitimate to extend the above writers' views by stating that not only a minus or plus self-picture or picture of another compensates itself in a percept on the other side of the scale, but an unrealistic self-in-relation-to-other structure is invariably accompanied by a compensating and equally unrealistic self-other structure. The realistically competent subject who perceives himself and loudly proclaims to others his status as an ant fearful of being trodden down by the surrounding gods invariably harbors the compensating, although possibly unconscious, image of himself as a proud god surrounded by negligible ants. Example is given by those who give up the very source of their psychic energy to images of significant others and yet may still enjoy the triumph of carrying the other about within one; of being the other and yet more than the other.

There is some disagreement by major theorists on this issue. We confess to presenting the subtleties of their arguments in crude form to emphasize the point at issue. Rogers believes that the neurotic is characterized by a

---

Freud then speaks of the interactions of the four structures:

"His hatred of his lady was inevitably coupled with his attachment to his father and inversely his hatred of his father with his attachment to his lady."

Freud concludes the paper as follows, p. 383: ". . . my patient . . . has disintegrated into three personalities: into one unconscious personality . . . and into two preconscious ones between which his consciousness could oscillate. His unconscious comprised those of his impulses which had been suppressed at an early age and which might be described as passionate and evil impulses. In his normal state he was kind, cheerful and sensible—an enlightened and superior kind of person, while in his third psychological organization he paid homage to superstition and asceticism. Thus he was able to have two different creeds and two different outlooks upon life. This second preconscious personality comprised chiefly the reaction-formations against his repressed wishes, and it was easy to see that it would have swallowed up the normal personality if the illness had lasted much longer. I (am) studying a lady suffering severely from obsessional acts. She has become similarly disintegrated into an easy-going and lively personality and into an exceedingly gloomy and ascetic one."

negative evaluation of himself. Horney states that the neurotic invariably has an idealized picture of the self. Indeed, according to her last major work (17), the perverted channeling of the "growth forces" of the "actual" individual into the idealized self-concept is the major pathogenic agent in the neurosis. Horney emphasizes the interaction of the real self and idealized self, although she earlier focused upon the neurotic's feeling of inadequacy and helplessness. Despite pointing out the negative self-picture as one of the chief motivating forces for the idealized self, she makes no mention of its being a necessary balance in the psychic economy. Sullivan (36) appears to assume that there is but a single self-percept.[11] The self is therefore either positively or negatively evaluated. Fromm believes that the neurotic represses the most positive parts of himself. We may infer that he regards the neurotic as having a negatively-toned conscious self-concept.

Thus, while Rogers believes that the patient's task in therapy is to attain an increasingly positive evaluation of the self, Horney stresses that the neurotic patient's task is to puncture the idealized image of the self. While Sullivan largely believes that the self is only what the individual is aware of, Fromm (14) follows Jung and Rank in believing that positive parts of the self-percept may be repressed. Both Rogers and Sullivan in their original studies of the self-structure seem to make contradictory formulations as to whether there can be a vital part of the personality that is dissociated from the total structure of the self.

Ferenczi (7) had long before correctly criticized Adler, whose position is similar in some ways to the later one of Horney's, because Adler (2, 3) was content to state that the drive for success follows from the feeling of inferiority. Ferenczi claimed that inferiority feelings themselves were often reactions to an exaggerated feeling of omnipotence to which such patients have become fixated in their early childhood. He viewed the seeking for greatness as an action to regain the omnipotence enjoyed in early childhood.

These differing views may all be seen to be partial views of our assumption that there is a multiplicity of self-in-relation-to-other percepts; that whenever an evaluation is out of line with an "objective" judgment, and the individual is intelligent enough to appreciate all the qualities of the object and is in direct contact with it, then the distorted evaluation invariably has a counterpart evaluation that goes in the other direction. The first and second evaluation support and compensate each other. Without the one, the other could not long maintain stability. This view necessitates a modification of what is conceived to take place in such defenses as denial, avoidance, i.e., restriction of the ego, repressions, etc. For it implies that the individual has no choice in denying vital data to which he is exposed and capable of understanding. His only choice is in how he would organize the data.

---

[11] In a work that appeared shortly before his death, Sullivan appeared to approach the view of a multiplicity of self-personifications in the individual personality (37).

The more general formulation proposed here may provide a framework for all the varying emphases of the different theorists. Rogers and Fromm have generalized from the conscious self-picture of certain neurotics to state that the neurotic always has a negative self-picture, while Horney in her last major work emphasized the neurotic's positive self-picture. Both views are partially true. For, as all these writers imply, the possession of one distorted evaluation is invariably accompanied by another. However, it should be noted that it does make a great difference as to whether the unconscious or the conscious picture is the negative self-percept. The first individual, with a conscious positive self-picture and an unconscious negative one is characterized by a false and brittle aggressiveness and cockiness. The second individual, with conscious negative picture and an unconscious positive one, is embarrassingly self-depreciatory, and yet intensely resentful when the other does not reassuringly speak of his omnipotence.

Ferenczi's view of the omnipotence of the child also seems a partial one in this context. Even the young child must be acutely aware of its inadequacies and its dependence upon strong significant others. Omnipotent feelings do not just come from feeling that one has all one wants and that there is nothing left to wish for, as Ferenczi appears to imply. An individual may reduce his expectations of the world to a minimum in order to reach for a cheap sense of omnipotence. Still the essential feeling of omnipotence is not that one's wants are temporarily sated, but rather that one permanently controls all the pertinent surrounding resources and the conditions to their access. The child early learns that he does not have such control.

Rogers has made the self-concept the focal point of his personality theory, and its change from a negative to a positive direction the focal point of his theory on psychotherapy. Let us therefore give his views further attention. This elaboration is further warranted because of the many experimental studies derived from this approach. Should our presentation be correct, we may see the limitations of the many excellent experimental studies done under Rogers dealing with the self-evaluation of the individual (28, 33, 35). They have found an increasingly positive self-evaluation reported by clients with successful therapy. It is true that there is verbalized an increasingly positive self-picture. In the context of the approach presented here, however, it can only mean, if the therapy is genuinely successful, that the client is integrating his previously unconscious overly-positive self-picture with the conscious negative one. Unless such integration takes place, the negative self-picture simply becomes unconscious. The positive self-percept then acts largely as a defense against the now unconscious self-percept. Further, it serves to ingratiate oneself with the counselor and congratulate him for his therapeutic efficacy by meeting his expectations of success—expectations generally based on the poor resolution by the counselor of his own narcissistic needs.

The non-directive technique according to Haimowitz (16), selects to itself and deals most adequately with neurotic intropunitive individuals whom we may assume have a low self-evaluation. It may readily be seen, however, that the task in therapy of a paranoid or manic patient with gradiose delusions is to decrease the positive self-evaluation that he reports at the beginning of therapy. The Stock and Scheerer studies, also done under Rogers, which revealed an increasingly positive evaluation of others in cases of successful therapy, may therefore be based upon a very selected population. In addition to assuming that the changes in the self-image are the independent variable from which ensue changes in the other-image, they tend to overlook the many cases of dependent individuals who rate others as too all-powerful or strong. These dependent individuals tend to either run away from or lean on others, depending in large part on whether the mighty other is seen as hostile or friendly. The task of these individuals in therapy is to decrease the conscious evaluation of others as all-mighty. The deflating of the other-image may be done by integrating the split pictures of the others' strength, an integration often accomplished by obtaining insight into one's unconscious depreciation of the others' capacity. It is realized, however, that some dependent individuals have a conscious depreciation of others and yet unconsciously regard them as supermen.

A maladjusted individual may begin therapy with any of the following largely *conscious* self-other evaluation constellations:

| Self | Others |
|---|---|
| 1. Overly positive (in relation to the objective qualities of the self) | 1. Overly positive |
| 2. Overly negative | 2. Overly positive |
| 3. Overly positive | 3. Overly negative |
| in addition to | |
| 4. Overly negative | 4. Overly negative |

which last alternative is the only constellation allowed for by the Rogerians.

It will be seen that the task in therapy for a subject having any one constellation is somewhat different from the task of an individual who may be placed in another category.

Once a split picture has occurred, it is very easy for it to continue in existence. The pathogenic structure often encourages the continuance of pathogenic dynamics. Let us assume that an individual performs acts that range with normal curve frequency in adequacy from a low of 2 to a high of 9. If two split self-images are largely organized about his adequacy, he may have the first self-picture, that of the self as inferior, assimilate all actions that may be valued at 2-5, while the other self-picture correspondingly assimilates to itself actions 6-9. The self-picture as one who is inferior is never modified by an action on the 6-9 area of the scale as this picture is never in figure when the person is organizing this positive data about himself. Similarly, the positive self-picture is dormant when a little-valued act occurs. Only that self-picture is dominant which "resonates" with the value of the act that is being assimilated to the self. (Indeed, that self-picture is largely responsible for the act.) Thus the compensating image is necessary in order to handle acts that would otherwise call for reorganization of the first image. Split images are in a sense "cheap" ways of avoiding the effort at the maximal structure of an integrated or differentiated perception.

Reactive social factors contribute their share to the growth and maintenance of these split structures. For example, if one acts out one's unwarranted belief of being a lord amidst lackeys, one will be rebuffed. Such rebuff will tend to reinforce the perception of the world as rejecting and aloof, and the self as inferior. In vicious cycle, new impetus is given to the compensatory percept of the self as possessing lordly qualities. Many relatively inadequate girls of the lower classes, unable to find a willing mate in their own class, will voice a very unrealistic desire for an outstanding member of the middle or upper class. An overly positive image of the self is used to cover the image of the self as inadequate and to depreciate the immediate outside world as most unsuitable. This process permits the girl to avoid doing any courtship battle, to chance being rejected, and having to revise her self-other structures. Balanced split structures, as do all self-other structures, seek out other individuals who very kindly aid in supporting the structure. Often it appears that the members of social groups are in conspiracy to succor each other's unrealistic self-other structures.

Once a split self-other structure has arisen, we may predict the interaction of the two structures by knowledge of their content. As Horney has pointed out, the idealized self-concept must hate the individual's actual being. All the more intense must be the hatred of the compensating structure, the negative self-concept, which is according to this theory of split and compensating structures, the symbol about which the rest of the individual's actual being must be organized.

## K. MEANS OF MITIGATING CONFLICT BETWEEN SELF-OTHER STRUCTURES

Even when to the observer conflict between two or more self-other structures in a personality seems inevitable, there are many possible avenues by which such conflict may be avoided or mitigated. It may be seen that some of these ways are healthier than others and that many overlap.

Two structures may be organized in the personality in a time-sequential manner which minimizes conflict. Two self-other structures may have logically contradictory themes, yet psychologically, the expression of one structure may be necessary to free the other for expression. If the two self-other structures are always dominant at different times and in different situations, they may be easily insulated from each other.

Insight may play a mediating rôle in decreasing conflict between self-other structures. The individual[12] may rationally integrate the assumptions underlying each self-other structure so that each assumption is a derivative of a major set of values, or at least stems from a non-contradictory set of assumptions. By perceiving any logical relationships of his different self-other structures, the individual may ease any conflictful aspects. By insight into the necessity of expressing an unpleasant or sacrificial self-other schemata in behavior, the individual may be more willing to accept a self-other structure that does not readily jibe with other structures in the personality. The individual may integrate via insight split and balancing self-other structures. The emergent structure would then accord more with reality. If the individual learns that the sequential acting out of two self-other structures is accompanied by the sequence of pain-pleasure, he will see the present painful self-other structure in its total context. When having to act an unpleasant transitional self-other structure in order to attain a position enabling one to act out a goal self-other structure, the individual may anticipate the goal-structure. Becoming a gatekeeper of the goal rôle and making others suffer in the same manner as he now suffers may be included as part of his pleasant imagery, giving support to playing the means-rôle. An individual may accept as wholesome the principle of multiplicity of self-other structures. He may, for example, pride himself on his ability to relate to many different categories of people.

---

[12] The term "individual" used throughout this section as one who does things to his self-other structures simply refers to another self-percept of the personality or a self-percept emerging from the constellation of existent ones, or to the constellaion of self-structures. Note again that each of these self-percepts is modified by the context provided by its respective significant-other representation and the representation of the interrelationship.

The absence of conscious insight may also help to eliminate conflict. The individual may lessen conscious conflict by employing unconscious defensive mechanisms against a conflictful self-other structure. For example, the individual may repress a feared self-other schema. Specious logical connections between self-other themes may also alleviate tension, at least on a conscious level.

Overt action may aid the individual to decrease tension between self-other structures. The individual may create or provoke the situation to which he must respond. Thus the individual may withdraw to a sub-culture in which he can act out consistent self-other structures. The individual may ally himself with others and change the group structure in a direction that would facilitate the expression of given self-other schemata without their conflicting or at least conflicting much less with other self-other structures. The feminsist movements facilitated woman's ability to take a more active and independent rôle. As a corollary to the above, the individual may remain aloof from any situation that threatens to put conflicting elements simultaneously in figure.

Other methods of diminishing conflict may utilize both action and fantasy. One may distort the expression of a legitimate and accepted self-other structure to give expression to a repressed self-other theme. The martyr may ensure getting himself killed in such a way as to gratify his fantasies of being homosexually attacked. There may be a convergence of logically contradictory themes into one action. Frequently, the more themes that converge upon one act, the more gratification does the act give. The individual acting out a given self-other theme may blur the line between self and other sufficiently to identify with and derive some of the pleasures that accompany the complementary self-other rôle of the other with whom he is interacting. Thus the martyr who has repressed the self-other fantasy of his destroying others may give this structure expression by identifying with his oppressors. The individual may mould his self-other schemata to fit into the expectations of significant others. He may, for example, dare to consciously evaluate others only as they insist upon being evaluated. The individual may act out one self-other structure in reality and a contradictory one in fantasy. The fantasy may be regarded by the person as legitimate compensation for the sacrifice to reality.

## L. Positive Effects of an Ideology of Multiplicity

Among the ways of reducing conflict listed earlier is the acceptance of multiplicity of self-other structures as a principle determining personality organization. Much of the shame before others and the guilt within oneself are caused not by the self-other structures that lie deeply buried in the self, but rather by other surface structures. Shame, for example, may be caused not only by the experience: '"What would he think of me if he knew my most secret dynamics that are even largely hidden from me!" In addition a shame-producing theme is: "What would he think of me if he knew how differently I behave with my wife or kids!" How reconcile in one's mind the difference between familial and occupational behavior is an unconscious problem that may loom as large for some as repressed incestuous desires. The supposed support that many wives give to husbands browbeaten by jobs is robbed by the husband's confusion in being looked up to by one, and down at by others.

The mobile status of many individuals in our culture frequently results in multiple self-other structures. Unfortunately, these tend to be guiltily isolated from each other because of the belief in the homogeneity of the ideally moral as well as healthy individual. Were multiplicity and even contrast accepted principles governing personality structure, much of the negative emotion accompanying a necessary multiplicity might be eliminated. Such an ideology commending multiplicity would tend to result in flexibility and spontaneity. An ideology of multiplicity would lead to the permissive interaction of self-other structures, giving rise to emergent and exciting perspective of self and other. With such an ideology would come greater awareness of the risks of multiplicity, and measures to preen and control excessive offshoots of the personality. The mature individual would recognize and be able to accept that he is necessarily *a* in situation *a, b* in situation *b,* etc.

Instead of this manner of fruitful interaction, we often find a dissociation of self-other structures in the futile effort to attain integrity. Unfortunately, this effort often leads to the very selling out of self that one sought to avoid. All too often potentially valuable self-other structures remain at a primitive level by virtue of their relegation to an unconscious status. The unwitting sacrifice of the richness of personality is often accompanied by the attribution of undeveloped self-other themes to the interaction of others and a passive intrigue with spiritless heroes of the culture.[13] There is the surrender of the opportunity to look at one's total thematic organization and to search for superior ways in which the entire constellation might be better organized. Perhaps only in this way can individuals hope to achieve a syntactic rather than a parataxic relation to the world.

---

[13] Avid sport fans frequently exemplify this process. Athletic events, as they are personalized conflictful situations, are ideal targets for such projections.

The personality type of controlled and rational multiplicity may well grow only in a society which encourages this manner of personality structure. It may also be the ideal type of personality for a civilized democratic society. But only if such a personality structure is idealogically approved, may the individual extend himself in different ways without suffering conflict as a result. It is indeed ironic that intellectuals who applaud a pluralistic social structure should have swallowed the bromide of the "integrated-personality."

## M. DISCUSSION

The approach presented here, though emphasizing the multiplicity of the normal ego structure, contrasts with that of the behaviorists who regard personality as a mere sum of stimulus-response elements. As the tested organism is placed by them in a situation which produces relatively isolated functioning of the given part, behaviorists experience no need to examine either other segments of the personality or their interrelations.

We differ with the many theorists, e.g., Adlerian and Sullivanian deviants from Freudian psychology, who bitterly contest the seeming fragmentation of the personality by Freudians into the mental institutions of superego, ego, and id with cross-cutting conscious, pre-conscious, and unconscious regions. Even Freudians (15) have been apologetic about the apparent reifications that resulted from this manifold division of the psyche. We have here pleaded cause not only for acceptance of the concept of different depth levels, but also for that of perpendicular-to-the-surface divisions as being characteristic of the normal organism, i.e., for differentiation of self-other structures along the surface of the personality facing the world. The differentiation results in varying activity by the given themes in the organism's different situations.

Although the human organism is a whole, it is a complex one in which many parts do not have conscious communicative accessibility to other parts. As some wholes are tightly organized, so are others very loosely tied together. It has *not* been our purpose to deny any interdependence of the sub-structures of the organism. In noting this interdependence, we stress however, as corollary to the Gestalt thesis that one must understand the whole to understand the part, that one must also comprehend the parts to predict interrelations. For example, one determinant of the manner of interraction between neighboring structures of the personality, and thus of the nature of the barrier conceptualized as separating them, is the thematic content of the structures. Should one structure define the self as having strong sexual desires and the other structure view sex as evil and the self as asexual, we may predict that in times of overt quiesence the barrier between the two will be very rigid. Further, when the barrier is permeated, the individual will experience and display intense disturbance.

The postulate of multiplicity of the normal ego structure makes necessary more than token reference to such structural dimensions, frequently offered by Gestaltists, as surface-depth, peripheral-central, and figure-ground to characterize the current state of a self-other structure. Two other dimensions which must be used to characterize every structure are conscious-unconscious, and extensive-non-extensive, i.e., the percentage of time a given structure tends to be in figure. For further understanding of a self-other structure, one should also know the content of its surrounding regions and its manner of interaction with them.

The Gestalt concept of integration remains an appropriate one to refer to both the arrangement of subordinate themes by a dominant one when the latter is activated and to the ordering by a given self-other structure of the resources of the organism, e.g., skills and talents, so as to best aid it to actualize itself.

The description of conflicting forces promising disturbance or having already disturbed an ego struggling to retain or re-establish its "integration" or equilibrium may best be conceptualized as either of the following:

1. The permeation of one structure by another, particularly the invasion of a defending and conscious schema by the defended-against and unconscious one.

2. The trauma of environmental stimuli that threaten the existent total structure of conscious schemata either by being congruent with, and thus activating a defended-against schema, or by the absence of any theme that can assimilate the novel stimuli.

Unity theorists correctly argue the need of the organism to identify itself in order to recognize its own interests as opposed to those of others. Still this distinction does not necessarily mean that the sense of selfhood cannot be distributed among a number of self-structures. This distribution requires further decision by the active self-other themes as to the legitimacy of interests of the different self-structures.

Gestaltists, e.g., Koffka and Lewin, and Gestalt-oriented theorists, e.g., Federn[14] and Rogers, in claiming that the normal individual has only one self-percept have perhaps been confused by the sense of myness that belongs to each conscious self-representation as it takes its turn at being in figure. Still, the different self-representations need only share the sense of myness. The remainder of the experience of the self may vary widely in different situations.

---

[14] Readers may note the many references to K. Lewin in this psychoanalyst's work.

The sense of myness occasionally may apply to a number of personality themes at one time. A sense of constancy and continuity with the past contributes to the sense of selfhood. A self-structure possesses a sense of continuity with the history of its past relationships to other self-definitions. A sense of communality is shared with other self structures from which a given theme has originally differentiated itself or with which it was temporarily joined or even shared the same fate, e.g., suppression. A given schema also obtains the experience of continuity in the stable patterned relationships to other schemata, permitting an emergent sense of myness to simultaneously apply to a number of schemata. Each self-other structure when in figure will tend to have a constancy of "perspective" as it peruses its relatively stable neighboring personality regions. As a result of this perusal process, the particular self-definition will be further differentiated by contrasting itself with other self-definitions as well as by obtaining a sense of sharedness, i.e., "weness," with the other self representations. For example, an individual splashing about in the swimming pool after a workout may well observe, "I'm a poor swimmer and maybe afraid of water. On a basketball court, however, I'm a superior and confident athlete." We suggest, nevertheless, that it is virtually impossible for anyone to be consciously aware of all his ego structures within a limited time period. Also contributing to the permeation beyond a given self-structure of the sense of myness are the many shared standard reference points of some self structures, e.g., one's features, one's clothing, one's physical environment, etc. However, it should be cautioned that even these constancies may be perceived somewhat differently with the activation of a different self-definition.

In a sense, the self-other structures are partners sharing the same vehicle. Due to the partially polemical nature of this paper, we perhaps have failed to stress that in the healthy individual each theme which has access to the motoric system, is somewhat aware of the whole[15] and will modify its behavior in the interests of other structures. For if a given structure acts itself out, the whole organism, including the other structures, may be held responsible. For example, a self-other structure may state that the best way to deal with exploitative authority figures—and that there are no other kind—is aggressively. If this theme realizes itself by punching the traffic cop on the nose, the remaining self-other structures, e.g., those oriented about sex and autonomy, are forced to suffer deprivation. It may be noted that all these ways in which a sense of myness may extend beyond one self structure are independent of any logical interlocking of the different self-definitions.

---

[15] Phenomenally, the whole is often vaguely recognized by the self-concept in figure and has frequently a distinct quality of *otherness* that accompanies its restrictiveness.

## N. Summary

The ego as subject synthesizes and organizes the symbolic processes of the organism according to the principles provided by the different assumptions regarding self and given other and the interrelationship. Thus the ego as subject is defined as identical with the ego as object.

The experiences peculiar to the different situations of the individual tend to become organized in stable and somewhat isolated assumptions about the self's qualities, the qualities of the significant others, and the interpersonal relations of self and other. The particular self-other structure then resonates with its corresponding situation, i.e., each draws the other forth. Each self-other representation tends to seek expression in behavior in the respective situation. The behavior, including that of the "provoked" significant other, is then assimilated to the given self-other schema, resulting in its reenforcement or revision. There is a continual and cyclic moulding and provoking of one's environment as determined by each self-other schema, followed by a review of the interaction, with a consequent strengthening, or a further differentiating and modifying, or more rarely, by a basic overhauling of the given theme. As there are a multiplicity of self-other representations, and as the ego is defined here as the constellation of self-other themes in the personality, the ego may best be conceptualized as a multiple structure. Indeed, maintenance of this multiple state is one of its primary goals.

## References

1. Abraham, K. Selected Papers on Psychoanalysis. London: Hogarth, 1927.
2. Adler, A. The Neurotic Constitution. New York: Moffat, Yard, 1917.
3. ———. Understanding Human Nature. New York: Greenberg, 1927.
4. Allport, G. W. The ego in contemporary psychology. *Psychol. Rev.*, 1943, **50**, 451-78.
5. Angyal, A. Foundations for a Science of Personality. New York: Commonwealth, 1941.
6. Federn, P. Ego Psychology and the Psychoses. New York: Basic Books, 1952.
7. Ferenczi, S. Sex in Psychoanalysis. Boston: Badger, 1916.
8. Freud, A. The Ego and the Mechanisms of Defense. London: Hogarth, 1937.
9. Freud, S. Three Contributions to the Theory of Sex. *Basic Writings of Sigmund Freud.* New York: Random House, 1938.
10. ———. The Ego and the Id. London: Hogarth, 1927.

11. ———. Formulations Regarding the Two Principles in Mental Functioning. *Collected Papers* (Vol. 4). London: Hogarth, 1925.
12. ———. A Case of Obsessional Neurosis. *Collected Papers* (Vol. 3). London: Hogarth, 1925.
13. ———. Contributions to the Psychology of Love. *Collected Papers* (Vol. 4). London: Hogarth, 1925.
14. FROMM, E. The Forgotten Language. New York: Rinehart, 1951.
15. GLOVER, E. Grades of ego differentiation. *Inter. J. Psychoanal.*, 1930, **11**, 1-11.
16. HAIMOWITZ, N. R. An Investigation Into Some Personality Changes Occurring in Individuals Undergoing Client-Centered Therapy. Unpublished Ph.D. dissertation. Univ. Chicago, 1948.
17. HORNEY, K. Neurosis and Human Growth. New York: Norton, 1950.
18. ———. Our Inner Conflicts. New York: Norton, 1945.
19. JAMES, W. Psychology. New York: Holt, 1890.
20. JUNG, C. The Integration of the Personality. New York: Farrar & Rinehart, 1939.
21. LOGAN, R. W. Book review. *The Nation,* 1951, **173**, 427-428.
22. *The Nation,* 1952, **174**, 388.
23. PERLS, F. Ego, Hunger, and Aggression. London: Allen & Unwin, 1947.
24. KOFFKA, K. Principles of Gestalt Psychology. New York: Harcourt, Brace, 1935.
25. LEWIN, K. A Dynamic Theory of Personality. New York: McGraw-Hill, 1936.
26. ———. Principles of Topological Psychology. New York: McGraw-Hill, 1936.
27. MURPHY, G. Personality: A Biosocial Approach to Origins and Structure. New York: Harper, 1947.
28. RAIMY, V. C. Self-reference in counseling interviews. *J. Consult. Psychol.*, 1948, **12**, 153-63.
29. RANK, O. Will Therapy and Truth and Reality. New York: Knopf, 1947.
30. RIESMAN, D. The Lonely Crowd. New Haven: Yale Univ. Press, 1950.
31. ROGERS, C. Client-Centered Therapy. Boston: Houghton-Mifflin, 1951.
32. ROSENMAN, S. Some unresolved problems in ego psychology. *J. Gen. Psychol.* (to be published).
33. SHEERER, E. T. An analysis of the relationship between acceptance of and respect for self and acceptance of and respect for others in ten counseling cases. *J. Consult. Psychol.*, 1949, **13**, 169-175.
34. STERBA, R. Fate of the ego in analytic therapy. *Internat. J. Psychoanal.*, 1934, **15**, 117-126.
35. STOCK, D. An investigation into the interrelations between the self-concept and feelings directed toward other persons and groups. *J. Consult. Psychol.*, 1949, **13**, 176-180.
36. SULLIVAN, H. S. Conceptions of Modern Psychiatry. Washington, D. C.: White Foundation, 1945.
37. ———. Multidisciplined Coördination of Interpersonal Data. In Sargent, S. S., and Smith, M. W. (*Eds.*), *Culture and Personality*. New York: Viking, 1951.
38. SUTHERLAND, E. H., & LOCKE, H. J. Twenty Thousand Homeless Men. Chicago: Chicago Univ. Press, 1936.

*Division of Testing and Guidance*
*The City College*
*Convent Avenue & 139th Street*
*New York City 31*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 421-426.

# ANXIETY LEVEL, RESPONSE ALTERNATION, AND PERFORMANCE IN SERIAL LEARNING*

*Department of Psychology, University of Pittsburgh*

J. B. HUGHES, II, J. L. SPRAGUE, AND A. W. BENDIG

## A. INTRODUCTION

Several recent articles have reported comparisons of the verbal serial learning of "anxious" and "nonanxious" *S*s using the anticipation method. The results of these studies are somewhat contradictory: Taylor and Spence (11) found nonanxious *S*s learned more rapidly; Montague (9) reports differences between groups to be a function of the association value of the stimuli with anxious *S*s learning faster than nonanxious with stimuli of high association value, with the reverse being true for low association stimuli; while Deese, Lazarus, and Keenan (1) found anxious *S*s generally superior to nonanxious *S*s in one sample and the reverse in a second sample. Such results may be a function of the intratrial and intertrial time intervals used with the anticipation method. Apparently there are differences among these studies in the intervals used which Hovland (5, 6) has shown to affect the rate of learning.

A second experimental variable that may be responsible for these differences is the type of learning task used (9). The learning task used by Taylor and Spence in their pioneer study (11, 12) was a sequence of 20 words presented to *S*s by a memory drum. Only two stimulus words were used, "right" and "left" which were presented to the *S*s an equal number of times in a randomized sequence. The *S*s were required to anticipate the stimuli on this list and learning was considered achieved when the criterion of two errorless trials was reached. The present authors noted that the particular sequence of "right" and "left" responses given by the *S*s required 12 alternations from a "right" to a "left" response, or vice versa. Such an alternation sequence would seem to favor the anxious *S*s on the basis of a decreased development of inhibitory potential ($I_R$ and $_sI_R$) for this task when compared with a learning task requiring less alternation of response (7). The present study was designed to replicate that by Taylor and Spence, using a different intratrial time interval and stimulus lists with both greater and lesser alternation of response within the lists.

---

*Received in the Editorial Office on June 20, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

## B. PROCEDURE

### 1. *Subjects*

The *S*s were selected from our introductory psychology course on the basis of extreme scores on the Manifest Anxiety Scale developed by Taylor (10). None of the *S*s had previously participated in any psychological investigation and they were informed that they had been selected as a representative sample of all college students.

The anxious group consisted of 27 *S*s with *MAS* scores of 25 or above and the nonanxious group of 27 *S*s obtained scores of 11 or below on a shortened 50-item form of the *MAS* (8). These two scores fell at the 37th and 89th percentile points for a group of 1971 students at the State University of Iowa and duplicate the cut-off scores used by Taylor and Spence in their original study. The *S*s in the two groups were group matched on the basis of sex and age, and were randomly assigned to the three learning list tasks described below.

### 2. *Apparatus*

A Hull-type memory drum, mounted on a table behind a screen, was used for presenting the lists of stimulus words. The choices were presented serially with a 2-sec. exposure time and a 2-sec. blank interval between exposures. A black shutter was dropped over the viewing slot during the 2-sec. interstimulus period so that *S*s did not observe the changing of the words on the list. Thus the interstimulus interval was 4-sec. while the intertrial interval was 6-sec. Apparently Taylor and Spence (11) omitted the blank 2-sec. interval between successive stimuli resulting in a 2-sec. inter-stimulus interval, and required their *S*s to anticipate the next stimulus while viewing the preceding stimulus (12).

The learning task consisted of one of a series of 20 stimulus words typed on a continuous white tape. Each choice consisted of either the word "right" or the word "left" and an equal number of each word appeared on each list. The word "right," when it appeared, was typed on the right side of the tape, while "left" always appeared on the left side of the viewing slot. Three sequences of stimulus words were used: the original (0) list used by Taylor and Spence, LLRLRRLLRLRRRLRLLRRL, a new list (I), LLRRRLL-RLRRLLLRRLRRL, and the second new list (II), LLLRRLLLLLRRR-LRRRLRR. Note should be taken that while all three lists required an equal number of *R* and *L* responses, List 0 required 12 alternations of *R* and *L* responses, List I required 10 such alternations, and List II only 7 alternations of response.

The instructions to *S*s were of the usual type for serial learning and similar to those used by Taylor and Spence (11, p. 63). The *S*s were required to give an anticipatory response to each exposure except those on the first trial. Since responses on the first trial would be on a random basis they were not recorded and do not appear in the later statistical analysis of the data. The *S*s were instructed to anticipate the next stimulus word while looking at the black shutter covering the viewing slot between successive exposures. Thus the mean time interval between *S*'s anticipatory response and the appearance of the anticipated word in the viewing slot was approximately 1 sec. Trials were continued until each *S* reached the criterion of two successive errorless trials (11, p. 63).

In essence, the experiment was designed as a 2 x 3 factorial (two anxiety and three word lists) with nine *S*s in each of the six groups.

## C. RESULTS

Since the distribution of trial scores was significantly skewed, a square root transformation (2, pp. 199-202) was applied to the scores which resulted in an approximately normal distribution. The analysis of variance performed on these transformed trial scores can be found in Table 1. Differences between the three lists were significant at the .05 point, but the difference between the anxiety groups and the lists by groups interaction were not significant. The six variances comprising the error term were found to be homogeneous by Bartlett's test (2, pp. 195-197).

TABLE 1

ANALYSIS OF VARIANCE OF TRANSFORMED TRIAL SCORES

| Source of variation | Sum of squares | *df* | Mean square | *F* |
|---|---|---|---|---|
| Total | 51.467 | 53 | | |
| Groups | 0.084 | 1 | 0.084 | |
| Lists | 7.485 | 2 | 3.742 | 4.103* |
| Interaction: G x L | 0.127 | 2 | 0.064 | |
| Within | 43.771 | 48 | 0.912 | |

*Significant at the .05 confidence point.

The total error scores were more normally distributed and an analysis of variance was performed on the untransformed error scores. We hypothesized that the anxiety groups might differ in the shape of the learning function, so the error scores were Vincentized to check this possibility. The number of errors made by each *S* for each tenth of his learning curve was found and the mean number of errors per tenth determined for each group. The resulting analysis can be found in Table 2. Differences between Vincent trials were significant at the .01 point, and differences between lists, as in the trials analysis, were significant at the .05 point. The difference between the anxiety groups and all of the interaction mean squares were again found not to be significant. These results indicate that, while there was a significant reduction in error scores between the tenths, the learning curves were the same for both anxious and nonanxious *S*s. Table 3 gives the means of the groups on error and untransformed trial scores. Comparison with similar data reported by Taylor and Spence indicates that our *S*s learned the tasks much more quickly even when the omission of errors made on the first trial from our computations is taken into account.

TABLE 2
ANALYSIS OF VARIANCE OF VINCENTIZED ERROR SCORES

| Source of variation | Sum of squares | *df* | Mean square | *F* |
|---|---|---|---|---|
| Total | 10084.66 | 539 | | |
| Between Subjects | 5504.68 | 53 | | |
| Groups | 1.56 | 1 | 1.56 | |
| Lists | 771.13 | 2 | 385.56 | 3.92* |
| G x L | 13.79 | 2 | 6.90 | |
| *S*s within Groups | 4718.20 | 48 | 98.30 | |
| Within Subjects | 4579.98 | 486 | | |
| Trials | 1748.83 | 9 | 194.31 | 32.77** |
| T x G | 92.63 | 9 | 10.29 | 1.74 |
| T x L | 123.89 | 18 | 6.88 | 1.16 |
| T x G x L | 51.43 | 18 | 2.86 | |
| Residual | 2563.20 | 432 | 5.93 | |

*Significant at the .05 point.
**Significant at the .01 point.

TABLE 3
MEAN ERROR AND TRIAL SCORES FOR EACH OF THE SIX GROUPS

| | Errors | | | | Trials | | | |
|---|---|---|---|---|---|---|---|---|
| | Anxious | | Nonanxious | | Anxious | | Nonanxious | |
| List | Mean | *SD* | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| 0 | 52.0 | 19.9 | 55.4 | 22.1 | 12.9 | 5.5 | 14.3 | 6.4 |
| I | 61.0 | 44.2 | 57.6 | 37.1 | 15.0 | 7.3 | 15.2 | 8.7 |
| II | 30.0 | 21.1 | 33.2 | 24.2 | 8.6 | 4.5 | 9.2 | 4.5 |

Since two of the authors (JBH and JLS) had tested all of the *S*s, a between-experimenters variable was inserted into the analysis given in Tables 1 and 2, and in neither case was this variable significant, nor were the interactions of *E* and any of the other main variables significant. Total *ACE* scores were available for 42 of the 54 *S*s and an analysis of variance, similar to that in Table 1, was computed on these scores. In this analysis none of the differences between lists, anxiety groups, or the interaction of lists and groups were at all significant.

Taylor and Spence (11) and Farber and Spence (3) reported significant positive correlation between the total number of errors (for combined anxious and nonanxious groups) at each choice-point and the differences between mean number of errors made by the nonanxious and anxious groups at these choice-points. A similar analysis on our data yielded the following rank-difference correlations: List 0, —.13; List I, —.10; and List II, —.12. It is obvious that none of these correlations are significant and, indeed, are opposite in sign to those previously reported.

## D. Discussion

The use of a longer intratrial interval with the serial anticipation method resulted, as expected from Hovland's study, in faster learning of the task by our *S*s. This reduced task difficulty may be responsible for the complete lack of significant differences between the learning rates of our anxious and nonanxious *S*s. Apparently anxious-nonanxious differences in learning are partially a function of temporal intervals within the learning task used. However, use of a less difficult task does not fully explain our results when the overall difficulty level of the 20 stimuli was correlated with the difference in mean number of errors made by anxious and nonanxious *S*s for each stimulus. Previous research has shown consistent, significant, positive correlations between these two variables, while our results gave consistent, nonsignificant, *negative* correlations.

Unfortunately our second hypothesis predicting that the difference in learning between anxious and nonanxious *S*s would be a function of the intratrial response alternation could not be adequately tested since a necessary condition for such a test would be the appearance of differences between the two classes of *S*s. Inspection of the means in Table 3 offers little support for our hypothesis of a greater decrement in the performance of anxious *S*s with decreased alternation of response.

The major conclusion arising from this study is that differences in learning between extreme anxiety groups are apparently a function of the intratrial temporal interval used, although other hypotheses to explain our negative results, such as the sampling differences found by Deese, Lazarus, and Keenan (1), are possible.

## E. Summary

Subjects receiving extreme low and high scores on an anxiety scale were randomly assigned to the learning of one of the three lists of words. Lists were learned by 54 *S*s using the serial anticipation method with a 4-sec. intratrial in-

terval. The 20-word lists varied in the number of intratrial response alternations required; the number of alternations ranging from 7 to 12. Error and trial scores showed significant differences between the lists, but no differences between anxiety groups. Insignificant negative correlations were found between overall difficulty of single words and the difference between mean errors made by anxious and nonanxious *S*s. It was concluded that differences in learning rates of anxious and nonanxious *S*s is a function of the intratrial temporal interval used with the learning task.

## REFERENCES

1. DEESE, J., LAZARUS, R. S., & KEENAN, J. Anxiety, anxiety reduction, and stress in learning. *J. Exper. Psychol.,* 1953, **46**, 55-60.
2. EDWARDS, A. L. Experimental Design in Psychological Research. New York: Rinehart, 1950.
3. FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. Exper. Psychol.,* 1935, **45**, 120-125.
4. HILGARD, E. R., JONES, L. V., & KAPLAN, S. J. Conditioned discrimination as related to anxiety. *J. Exper. Psychol.,* 1951, **42**, 94-99.
5. HOVLAND, C. I. Experimental studies in rote-learning theory: II. Reminiscence with varying degrees of syllable presentation. *J. Exper. Psychol.,* 1938, **22**, 338-353.
6. ———. Experimental studies in rote-learning theory: III. Distribution of practice with varying speeds of syllable presentation. *J. Exper. Psychol.,* 1939, **23**, 172-190.
7. HULL, C. L. Essentials of Behavior. New Haven: Yale Univ. Press, 1951.
8. MCCREARY, J., & BENDIG, A. W. Comparison of two forms of the Manifest Anxiety Scale. *J. Consult. Psychol.,* 1954 (in press).
9. MONTAGUE, E. K. The rôle of anxiety in serial rote learning. *J. Exper. Psychol.,* 1953, **45**, 91-96.
10. TAYLOR, J. A. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.,* 1953, **48**, 285-290.
11. TAYLOR, J. A., & SPENCE, K. W. The relationship of anxiety level to performance in serial learning. *J. Exper. Psychol.,* 1952, **44**, 61-64.
12. ———. Erratum. *J. Exper. Psychol.,* 1953, **46**, 462.

*Department of Psychology*
*University of Pittsburgh*
*Pittsburgh 13, Pennsylvania*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 427-437.

# RORSCHACH SCORES IN THE MANIC STATE*

*Norwich State Hospital, Norwich, Connecticut*

HERMANN O. SCHMIDT AND CHARLES P. FONDA[1]

## A. INTRODUCTION

Some clinicians, confronted with a problem of differential diagnosis where one possibility is manic psychosis, may find themselves in a position similar to that related by Schafer (19), who complained that he had seen the test records of manics too infrequently to generalize about them. If the clinician has recourse to the literature in the field, he will find a diversity of opinion as to how Rorschach scores of manics should differ from those of normal persons. In actual practice, of course, the Rorschach is rarely needed to help distinguish manics from normals; more often, the problem is to determine whether a patient's disorder is the manifestation of a manic-depressive psychosis or one of the schizophrenic processes (9). (For a discussion of the Rorschach diagnosis of schizophrenia, see Kendig (10). Rorschach diagnosticians who feel that a subject's style of verbalization carries within it crucial indications as to the nature of the maladjustment, as Rapaport *et al.* (17) and the present authors do, will find in a paper by Lorenz (13) many useful comments about this aspect of manics' behavior.)

Rorschach (18) discussed the manic state, but his observations are based upon a total of only 14 patients of the manic and depressed groups taken together. "It is possible," he wrote (p. 71), "that my material does not represent all of the varied types of manic states." An extreme heterogeneity of symptom-patterns among manics has, indeed, been frequently noted (20, 21).

Levy and Beck (12) published in 1931 the scoring summaries of 11 manic (or hypomanic) psychotics, and Beck (1) later published the full protocols of four others. These authors state that their material is intended merely to illustrate the Rorschach findings in the psychosis in order to see what light they might throw on its psychopathology. They concluded from their investigation that the manic state increases *W*, *DW*, and *Sum C* but reduces *F*+%, *A*%, *FC*, *D*, and *Dd*. In summarizing this study, Bellak (4, p. 105) must have been the victim of a typographical error, for his text reads, "The

---

*Received in the Editorial Office on June 20, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Grateful acknowledgment is made to the Superintendent, Ronald H. Kettle, M.D., whose interest and coöperation made possible the publication of this paper.

manic phase is characterized by . . . a small number of 'whole' responses . . . and a high percentage of animal responses." The exact opposite, of course, is true. Levy and Beck considered two alternate possibilities regarding *M* responses in manics, both based upon rationale advanced by Rorschach (18). On the one hand, they say *M* should increase in the manic state because it indicates fantasy activity which derives its energy from affect; but, on the other hand, they would anticipate a decrease in *M* in the manic state because the augmented motility should drain off the *M* tendency.

By 1945, Beck (2) was still obliged to write that his own number of records of patients in manic excitement remained "small." Nothing further has appeared in the literature since that date.

## B. PROBLEM

The present investigation has for its purpose an analysis of the scoring summaries of Rorschach protocols obtained from all of the manic patients tested during the past seven years at a large state hospital. Such an analysis makes it possible to indicate: (*a*) how the scores of manics differ from those of normals; (*b*) how these compare with results obtained, or predictions made, by other investigators; and (*c*) how these differ from those of other psychotics, e.g., schizophrenics.

## C. METHOD

### 1. *Subjects*

*a. Manics.* From the files of the Psychological Laboratories at the Norwich State Hospital there were selected 42 Rorschach score summaries. These represent all of the *initial* protocols obtained during the past seven years from patients whose subsequent recorded diagnosis was manic-depressive psychosis (or manic-depressive reaction), manic type. It appears likely that this sample is minimally contaminated by nosological unreliability, since the manic-depressive diagnosis has been applied to less than three per cent of new admissions to this hospital, against a national rate of nearly eight per cent (4). The ratio of depressive to manic phase is on the order of two to one here. Many of the so-called "manic" cases described in the literature (5, 14, 15) would probably have received here a schizophrenic diagnosis because of the more rigid criteria employed at this hospital. Since the Rorschach protocols in the present study were obtained from testable patients, it is probable that the term "hypomanic" gives a more accurate description of their mental status at the time of testing. All were overtalkative and many were hyperactive; some were hostile and belligerent, while others gave an appearance of elation and euphoria. None was hallucinated, and while most of them elaborated delusional trends, the diagnosis was made on the basis of staff opinion that the psychotic behavior merely included, rather than emerged from, whatever bizarre notions might have been expressed by the patient. In order to be included in the present study, the Rorschach scores must have been obtained on the first test administered prior to diagnosis, during the patient's first month in the hospital, and before any specific treatment, such as *EST*, was begun. The test scores (Beck) were tabulated and means and sigmas computed; these were compared by critical ratios with the means derived from data on normal subjects published by Beck *et al.* (3).

*b. Schizophrenics.* From the same files, a second selection of Rorschach records produced 42 response summaries for protocols obtained during the past seven years from patients whose official diagnosis was dementia praecox (schizophrenia), or schizophrenic reaction. These records were selected in such a way that each schizophrenic could be matched with a manic patient on the basis of age, sex, and total number of Rorschach responses. Table 1 gives some descriptive characteristics of the three groups of subjects used in this study. Represented in the schizophrenic group were 15 paranoid, 15 catatonic, 3 hebephrenic, 1 simple, and 8 type-undetermined diagnoses. Test scores were tabulated and mean values computed for this group. In addition, the *differences* between scores for each schizophrenic and his matched

TABLE 1
DESCRIPTIVE CHARACTERISTICS OF THREE SAMPLES: MANICS, SCHIZOPHRENICS, AND BECK'C NORMALS

| | Manics (N.S.H.) ($N = 42$) | | Schizophrenics (N.S.H.) ($N = 42$) | | Normals (3) ($N = 157$) | |
|---|---|---|---|---|---|---|
| | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| Age | 41.9 | 11.6 | 41.1 | 10.5 | 30.5 | 8.6 |
| *IQ* (W-B)* | 104.5 | 13.7 | 100.9 | 14.5 | — | — |
| | % | | % | | % | |
| Sex | | | | | | |
| Male | 33.3 | | 33.3 | | 45.2 | |
| Female | 66.7 | | 66.7 | | 54.8 | |
| Education | | | | | | |
| College Grad. | 7.7 | | 2.5 | | 2.6 | |
| Some College | 18.0 | | 10.0 | | 10.4 | |
| H. S. Grad. | 25.6 | | 35.0 | | 35.1 | |
| Some H. S. | 18.0 | | 15.0 | | 39.6 | |
| Elementary Grad. | 12.8 | | 17.5 | | 10.4 | |
| Some Elementary | 18.0 | | 20.0 | | 1.9 | |

*Computed on 34 pairs for whom W-B *IQ*'s were available, $CR = 1.0$ (not significant).

manic "partner" were also tabulated. The means of these differences, together with their standard errors, were computed in order to estimate significance via critical ratios.

*c. "Textbook" manics.* From the 42 manic records, 10 were selected as most representative of the classical, or "textbook," clinical picture. Their scores were tabulated separately for comparison with: (*a*) those of the remaining 32 manics, and (*b*) the *initial* scores of 15 manics or hypomanics reported by Levy and Beck (12) and Beck (1).

### 2. *Hypotheses*

From the standard Rorschach manuals (1, 11, 16, 18) expectations or hypotheses regarding a variety of aspects of manics' scores were abstracted. Our own findings will be presented as possibly relevant tests of these several hypotheses.

## D. Results and Conclusions

Table 2 presents a comparison of manics with the matched group of schizophrenics and with a normal sample. Of the 50 (non-independent) differences evaluated in these comparisons, 17 reached significance at the .01 level and 2 at the .05 level of confidence.

A word of explanation regarding the data for normals in Table 2 is needed. Means and sigmas are from Beck *et al.'s* (3) Table 22, with the following exceptions: (*a*) Values for the *R* and *W* means were computed directly from their Table 4. (*b*) Mean values for $D+Dd$ were made equal to *R-W* and given the same proportions relative to each other as in Beck *et al.'s* Table 6. (*c*) Values for the *CF* and *FC* means were obtained by analysis of data in their Table 9 (p. 259), which gives an unweighted mean for "Pure color, including blends" together with weighted means for "Sum of color values" and for "*C* sum — *FC* sum"; from these data, computation of internally-consistent values for the unweighted *CF* and *FC* means was straightforward. (*d*) Values for *Hd* were computed directly from their Table 12. (*e*) Values for *TIR* were computed directly from their Table 14.

### 1. *Similarities*

Table 2 shows that there may be no real difference among the three groups with respect to over-all productivity (*R*), attention to the obvious (*D*), opposition trends (*S*), emotional lability (*CF*), or concern over health (*An*). The two *psychotic* groups probably do not differ in functional intelligence

TABLE 2

COMPARISON OF RORSCHACH SCORES FOR MANICS WITH THOSE FOR A MATCHED GROUP OF SCHIZOPHRENICS AND A NORMAL SAMPLE

| | Normal† ($N = 157$) | | | Manic ($N = 42$) | | | Schizophrenic ($N = 42$) |
|---|---|---|---|---|---|---|---|
| Item | Mean | *SD* | Diff. (Manic minus Normal) | Mean | *SD* | Diff. (Manic minus Schiz.) | Mean |
| *R* | 33.1 | 17.68 | 2.4 | 35.5 | 18.98 | — 0.1 | 35.6 |
| *W* | 5.9 | 3.76 | 2.2* | 8.1 | 4.45 | 1.6 | 6.5 |
| *DW* | | | | 0.3 | 0.83 | 0.0 | 0.3 |
| *D* | 24.1 | 10.49 | — 1.6 | 22.5 | 14.56 | — 0.9 | 23.4 |
| *Dd* | 3.2 | 3.38 | 1.4 | 4.6 | 6.76 | — 0.8 | 5.4 |
| *Po* | | | | 0.3 | 1.22 | 0.1 | 0.2 |
| *Z* | 22.5 | 14.91 | 13.7** | 36.2 | 23.79 | 9.0** | 27.2 |
| $\Sigma C$ | 3.1 | 2.74 | 3.3** | 6.4 | 4.95 | 1.4** | 5.0 |
| *C* | 0.6 | 0.92 | 1.3** | 1.9 | 1.83 | 0.4 | 1.5 |
| *CF* | 1.6 | 1.77 | 0.7 | 2.3 | 2.66 | 0.4 | 1.9 |
| *FC* | 1.2 | 1.21 | 1.3* | 2.5 | 4.11 | 1.0 | 1.5 |
| *M* | 3.5 | 3.24 | 0.0 | 3.5 | 2.85 | 1.2* | 2.3 |
| *Y* | 2.0 | 2.22 | 2.2** | 4.2 | 5.43 | — 0.5 | 4.7 |
| *V* | 1.8 | 2.05 | — 0.9** | 0.9 | 1.93 | 0.4 | 0.5 |
| *F*+% | 79.2 | 10.20 | —16.9** | 62.3 | 16.79 | 4.5 | 57.8 |
| *P* | 6.8 | 2.41 | — 1.1** | 5.7 | 2.33 | 0.6 | 5.1 |
| *A*% | 46.4 | 13.12 | — 6.3** | 40.1 | 10.40 | — 1.3 | 41.4 |
| *H* | 4.0 | 3.62 | — 0.2 | 3.8 | 2.82 | 1.6** | 2.2 |
| *Hd* | 2.3 | 3.10 | 0.2 | 2.5 | 3.94 | — 0.8 | 3.3 |
| *An* | 1.6 | 1.97 | 0.4 | 2.0 | 3.55 | 0.1 | 1.9 |
| *S* (all) | 1.9 | 2.14 | 0.9 | 2.8 | 3.09 | 0.4 | 2.4 |
| *S* (main) | 0.7 | 1.07 | 0.0 | 0.7 | 1.33 | 0.1 | 0.8 |
| *Lambda* | 70.2 | 13.71 | —11.6** | 58.6 | 14.06 | — 8.5** | 67.1 |
| *AR*‡ | 60.0 | 19.00 | — 0.8 | 59.2 | 26.27 | 2.0 | 57.2 |
| *T/IR* | 28.0 | 25.80 | —14.9** | 13.1 | 9.05 | —16.0** | 29.1 |
| *T/R* | | | | 37.6 | 17.17 | —16.7* | 54.3 |

*Significant beyond .05 level.
**Significant beyond .01 level.
†Data from Beck *et al.* (3). See text.
‡Affective ratio (R, VIII-X/I-VII).
Note.—*SD* for Schizophrenic group not computed. See text.

(*W*), alogical (*DW*) and illogical (*Po*) thought processes, or in primitive egocentricity (*Pure C*).

*The M response.* These manics seem not to differ from normals in the extent of their imaginative living (*M*). One possible explanation of this finding is that *both* of the processes assumed by Levy and Beck (12) exist simultaneously and cancel each other out. Another possibility for future exploration might be that *M* yields an index of degree of mania; i.e., in the hypomanic phase affective dilation increases the *M*, while the extreme motility of excitement may short-circuit this phantasy activity.

## 2. *Differences*

*a. Manics vs. normals.* Compared with normals, as shown in Table 2, manics respond more rapidly (*TIR*), show greater emotional dilation (*lambda*) and discharge more intellectual energy (*Z*). The manics' higher output of intellectual energy is reflected in their stronger tendency to generalize (*W*) and their less frequent resort to stereotypy (*A%*). That they are also psychotic is mirrored in the manics' inferior perceptual accuracy (*F+%*) and their inadequate participation in the conventional currents of thought (*P*). Their emotional dilation is reflected in higher scores for primitive egocentricity (*Pure C*) and available affective energy (*Sum C*); they also exceed the normals in capacity for feeling in tune with others (*FC*) as well as in feelings of dysphoria (*Y*), whereas they give considerably less evidence of capacity for critical self-appraisal (*V*). Many authors who discuss the motivational aspects of the manic state consider it a manifestation of the patient's attempt to contradict and deny the unhappiness and frustrations of his concurrent and seemingly insoluble anxiety situation (14). (The Pagliacci maneuver?)

*b. Manics vs. schizophrenics.* From the standpoint of differential diagnosis, Table 1 highlights several significant differences between the mean scores for manics and for the matched schizophrenics. This comparison shows that *people* appear more often in the content of the manics' responses. Thus, the manics give reliably more human movement (*M*) and whole human (*H*) responses than do the schizophrenics. (Lorenz (13) makes the cogent observation that one predominant mode of adaptation common to all manic patients is their constant attempt to dominate their *human* environment.) The manics also exceed the schizophrenics in affective dilation (*lambda*), in emotional energy (*Sum C*), and in intellectual energy (*Z*). Finally, the manics respond to the blots much more quickly than do their schizophrenic controls, both on the average (*TR*) and in the first response to each card (*TIR*).

## 3. *Trends*

If it be assumed that manic-depressives occupy a position on a continuum of pathology between normals and schizophrenics (7), then it might be instructive to note any trends which seem to accompany the postulated increase in malignancy, even though specific differences between pairs of values may not reach statistical significance. Thus, mean scores in the following categories are seen to grow in magnitude from normal through manic to schizophrenic: undirected attention (*Dd*), dysphoria (*Y*), and allocentric constriction (*Hd*); whereas the following means diminish: accuracy of reality-perception (*F*+%), conventionality (*P*), self-appraisal (*V*), affective contact (*R*, VIII-X/I-VII), and interest in others (*H*). Changes in the ratio of *H* to *Hd* are particularly interesting: for the normals and manics, this ratio is 4.0/2.3 and 3.8/2.5, respectively; among the schizophrenics, however, it is reversed, i.e., 2.3/3.3.

### 4. *"Textbook" Manics*

Means and sigmas for the scores of the 10 most typical manics in our group are given in Table 3 for comparison with similar data from protocols of the 15 cases reported by Levy and Beck (12) and Beck (1). The Norwich sample exceeds Levy and Beck's group in total number of responses (*R*) and, with the exception of *DW* and *A*%, in all other tabulated scores. For this reason, the scores in Table 3 are also expressed as percentages of *R* in order to facilitate comparison. Only in the cases of *DW* and *Y* do any of the percentage differences approach significance. This simply may reflect changes in the scoring criteria, however, inasmuch as these have undergone some alteration during the quarter of a century since the Levy and Beck records first were scored. A separate tabulation of scores for the remaining 32 patients in the Norwich sample yields a set of means (not shown) which differ reliably from those of the smaller group of 10 only in the case of *TIR* and *lambda*, both values being lower in the latter group.

TABLE 3
COMPARISON OF RORSCHACH SCORES FOR TWO GROUPS OF "TYPICAL" MANICS

| | Levy & Beck (12) ($N = 15$) | | | Norwich ($N = 10$) | | |
|---|---|---|---|---|---|---|
| Item | Mean | *SD* | % of *R* | Mean | *SD* | % of *R* |
| *R* | 28.00 | 13.06 | | 39.3 | 24.85 | |
| *W* | 6.87 | 3.32 | 24.54 | 9.8 | 6.16 | 24.93 |
| *DW* | 1.73 | 1.57 | 6.17 | 0.1 | 0.33 | 0.03 |
| $\Sigma C$ | 6.23 | 3.31 | | 9.2 | 7.25 | |
| *C* | 1.87 | 1.93 | 6.68 | 2.3 | 2.28 | 5.85 |
| *CF* | 2.40 | 2.00 | 8.57 | 3.7 | 3.21 | 9.41 |
| *FC* | 2.27 | 2.08 | 8.11 | 4.2 | 6.70 | 10.69 |
| *M* | 2.33 | 1.74 | 8.32 | 4.5 | 3.35 | 11.45 |
| $M/\Sigma C$ | 0.37 | | | 0.49 | | |
| *Y* | 1.27 | 1.57 | 4.54 | 7.5 | 8.63 | 19.08 |
| *Lambda* | 63.80 | | | 43.5 | | |
| *F*+% | 58.93 | 17.30 | | 62.2 | 17.58 | |
| *A*% | 40.93 | 11.77 | | 34.9 | 12.19 | |

### 5. *Hypotheses*

Table 4 shows the degree to which our 42 manics' records confirm the various hypotheses advanced by several authors. For example, Rorschach's predictions regarding *F*+% and *Pure C* receive confirmation, whereas the *normals* conform more often to his specifications for the manics in the case of *M, CF, FC, W,* and *A*%. His prediction concerning *Exp* is not substantiated. (Certain other hypotheses advanced by Rorschach will be considered below.) Beck's observations regarding *F*+%, *M, Sum C, Exp, P* and *S* are confirmed; but for *Pure C, CF, FC, A*%, *Z,* and *Y*, our manics gave results not entirely consistent with Beck's expectations. Klopfer and Kelley's statements are confirmed for *F*+%, Sum *C, Z,* and *S;* opposite results appear for *M, A*%, and *Y*. Phillips and Smith receive confirmation on *Sum C, A*% and *Y*, but most *M* or *CF*+*C*. .

*a. Rorschach's observations.* A number of additional hypotheses formulated by Rorschach (18), which could not be conveniently contained in Table 4, were also evaluated in the course of the present investigation, with the following results:

1. "In manic-depressive insanity," Rorschach wrote (p. 27), ". . . the better the form, the fewer the *M*'s." In our data the correlation between *F*+% and *M* is —0.14 (not significant, but in the predicted direction).

2. "There is a direct proportion between the number of *M* and *C* responses in . . . manic depressives" (p. 31). In our data, the correlation between *M* and *Sum C* is +0.05 (not significant, but in the predicted direction).

3. "Many . . . *M*— answers occur in protocols of manics . . . but are rare in schizophrenics" (p. 27). In our data 45.2 per cent of the manic patients gave *M*— responses as against only 26.2 per cent of the schizophrenics. The difference between these percentages is 1.79 times its standard error and reaches one-tail significance with a p/2 of .037.

4. "Reaction time in schizophrenics is shorter than in all other groups" (p. 22). In our data, *TIR* for the manics was, on the average, 16 seconds faster than for the schizophrenics. This difference is significant at the .01 level. It should be remarked in connection with this and the previous hypothesis, however, that our schizophrenics are probably not a representative sample, having been deliberately selected to match the manics.

*b. Use of white as color.* Our data permit the recording of one additional observation which may be relevant to Piotrowski's (6, p. 71) theory that "positive reaction to white as color should . . . indicate clearly conscious intermittent ecstatic moods marking the close and overcoming of a depressive

TABLE 4

PREDICTIONS BY VARIOUS AUTHORS CONCERNING RORSCHACH SCORE DIFFERENCES AND RESULTS OBTAINED BETWEEN 42 N.S.H. MANICS AND BECK'S 157 NORMALS

| | Predictions | | | | Results | | |
|---|---|---|---|---|---|---|---|
| Item | Rorschach (18, p. 51) | Beck (1, p. 55) | Klopfer & Kelley (11, p. 398) | Phillips & Smith (16, p. 84) | Mean Diff. (Manic minus Normals) | Percentage of cases within limits predicted Manics | Normals |
| $F+\%$ | 50-70 | Low | Drops | | —16.9** | 44** | 18 |
| $P$ | | Variable to low | | | — 1.1** | | |
| $W$ | 4- 7 | | | | | 22 | 31 |
| $S$ (all) | | High | Common | | 0.9 | | |
| $Z$ | | Variable to medium | Much (p. 397) | | 13.7** | | |
| $A\%$ | 50-70 | Medium | High | Low (p. 115) | — 6.3** | 17* | 36 |
| $C$ | 2- 6 | Medium | | | 1.3** | 46** | 6 |
| $CF$ | 2-3 | High | | | 0.7 | 15 | 19 |
| $FC$ | 1-3 | Medium | | | 1.3** | 19** | 50 |
| $\Sigma C$ | | Medium to high | Increase | 4.5 or more | 3.3** | 65** | 30 |
| $CF+C$ | | | | Generally exceeds $FC$ | | 63† | 65† |
| $Y$ | | Low | Not so prominent | Many | 2.2** | | |
| $M$ | 5 or more | Variable to medium | Increase | | 0.0 | 30 | 32 |
| $M$ | | | | 3-4 | | 14 | 12 |
| $Exp$‡ | Ambiequal (p.115) | | | | | 31 | ? |
| $Exp$‡ | | Extratensive | | | | 60 | ? |

*Note.*—See text for source of data for normals. Percentages of cases within predicated limits are estimated from parameters, assuming normal distributions, for each variable.

*Difference significant beyond .05 level.

**Difference significant beyond .01 level.

†Percentage of all color responses that are $CF$ and $C$.

‡*Exp* quantified as described in (8).

state." When we examined our patients' protocols for responses which could meet Piotrowski's criterion, we found them in four (10 per cent) of the schizophrenics' records and seven (17 per cent) of the manics'. The difference between these proportions is not significant ($CR = 0.97$).

## E. SUMMARY

The Rorschach scores of 42 manics were compared with those of 42 schizophrenics, matched on the basis of age, sex, and total number of responses. The manics' scores were higher than the schizophrenics' for *Sum C, Z, H,* and *M;* they were lower for *lambda, TIR,* and *TR.* Comparison of the manics with Beck's normals showed the manics to be higher on *Sum C, Z, W, Pure C, FC,* and *Y;* they were lower on *F+%, A%, lambda, P, V,* and *TIR.* A number of hypotheses advanced by various writers were evaluated in the light of these empirical findings. Three of 12 hypotheses advanced by Rorschach were confirmed, and a consistent trend was noted in two others. Also confirmed were six of Beck's 12, four of Klopfer and Kelley's seven, and three of Phillips and Smith's five.

## REFERENCES

1. BECK, S. J. Introduction to the Rorschach method. A manual of personality study. *Res. Monog. Amer. Orthopsychiat. Ass.*, 1937, No. 1.
2. ———. Rorschach's Test (Vol. 2), A Variety of Personality Pictures. New York: Grune & Stratton, 1945.
3. BECK, S. J., RABIN, A. I., THIESEN, W. G., MOLISH, H., & THETFORD, W. N. The normal personality as projected in the Rorschach test. *J. of Psychol.*, 1950, **30**, 241-298.
4. BELLAK, L. Manic-Depressive Psychosis. New York: Grune & Stratton, 1952.
5. BLALOCK, J. R. Psychology of the manic phase of the manic-depressive psychoses. *Psychiat. Quart.*, 1936, **10**, 262-344.
6. BRUSSEL, J. A., HITCH, K. S., & PIOTROWSKI, Z. A. A Rorschach Training Manual. Utica, N. Y.: State Hospitals Press, 1950.
7. EYSENCK, H. J. Schizothymia-cyclothymia as a dimension of personality: II. Experimental. *J. Pers.*, 1952, **20**, 345-384.
8. FONDA, C. P. The nature and meaning of the Rorschach white space response. *J. Abn. & Soc. Psychol.*, 1951, **46**, 367-377.
9. HOCH, P. H. Manic-depressive psychosis. *Med. Clin. N. Amer.*, 1948, **32**, 641-646.
10. KENDIG, I. V. Rorschach indications for the diagnosis of schizophrenia. *J. Proj. Tech.*, 1949, **13**, 1-8.
11. KLOPFER, B., & KELLEY, D. M. The Rorschach Technique. Yonkers, N. Y.: World Book, 1942.
12. LEVY, D. M., & BECK, S. J. The Rorschach test in manic-depressive psychosis. *Res. Publ. Ass. Nerv. Ment. Dis.*, 1930, **11**, 167-181. Also in *Amer. J. Orthopsychiat.*, 1934, **4**, 31-42.
13. LORENZ, M. Language behavior in manic patients. *AMA Arch. Neurol. & Psychiat.*, 1953, **69**, 14-26.

14. MASSERMAN, J. H. Psychodynamisms in manic-depressive psychoses. *Psychoanal. Rev.*, 1941, **28**, 466-478.

15. OLTMAN, J. E., & FRIEDMAN, S. The rôle of hostility in affective psychoses. *J. Nerv. & Ment. Dis.*, 1943, **97**, 170-196.

16. PHILLIPS, L., & SMITH, J. G. Rorschach Interpretation: Advanced Technique. New York: Grune & Stratton, 1953.

17. RAPAPORT, D., GILL, M., & SCHAFER, R. Diagnostic Psychological Testing: The theory, statistical evaluation, and diagnostic application of a battery of tests. Chicago, Ill.: Yearbook Publishers, 1946. (Vol. 2.)

18. RORSCHACH, H. Psychodiagnostics, a Diagnostic Test Based on Perception. (Translated by P. Lemkau & B. Kronenburg.) Bern: Huber, 1942.

19. SCHAFER, R. The Clinical Application of Psychological Tests. New York: Int. Univ. Press, 1948.

20. SKOTTOWE, I. A contribution to the psychopathology of mania. *J. Ment. Sci.*, 1936, **82**, 656-658.

21. WITTENBORN, J. R., & WEISS, W. Patients diagnosed manic depressive psychosis—manic state. *J. Consult. Psychol.*, 1952, **16**, 193-198.

*Norwich State Hospital*
*Norwich, Connecticut*

Published as a separate and in the *Journal of Psychology,* 1954, **38**, 439-452.

# THE RELATION OF SIZE OF LIGHT TO AUTOKINETIC EFFECT*

*Department of Psychology, McGill University*

ABRAHAM S. LUCHINS[1]

This is one of a series of reports on factors that influence the autokinetic phenomenon, the apparent movement of a stationary light in an otherwise dark room. The present paper describes exploratory experiments which varied the size of the light in an effort to determine the influence of size on the autokinetic effect.

## A. EXPERIMENT I

### 1. *Method*

The source of light was from a black painted wooden box 18¼ inches long, 19½ inches wide, and 11½ inches deep, with an opening 6¾ inches long by 4½ inches wide centered on its front panel. Inside the box was a bulb of 100 watts which was on throughout the experiment. Six slides, made of black, light-proof cardboard and with the same dimensions as the box's front panel, each had a square perforation in the center. One slide at a time was placed over the box's front panel, so that the light seen by *S* (the *A*-light or stimulus) corresponded in size to the size of the perforation. The lengths of the sides of the square perforations, in the order presented, were three-eighths, one-eighth, one, one-half, three-quarters, and one-quarter of an inch. The dark box was on a table 12 feet away from *S*'s chair with the front panel facing *S* and the *A*-light approximately at eye level.

Since the source of the light was held constant (the 100 watt bulb) and since the perforations varied in size, the total illumination also varied, with the larger perforations allowing more light to escape. It should be kept in mind that to this extent the factors of size and total illumination were confounded in this and in the subsequent experiments.

---

*Received in the Editorial Office on June 28, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1]This paper is a report on part of a project carried out under the auspices of the Defence Research Board, Ottawa, Canada. The assistants in this project, who were paid by the Defence Research Board, were Norman Endler, Rhoda Finkel, and Melvin Schwartzben.

Ten college students were tested individually. With his eyes closed, *S* was led into the room, seated in the chair, and had a shield placed over his eyes. The shield was used to prevent *S* from seeing the walls or the experimenter.

After five minutes *S* was asked to open his eyes, to look straight ahead (the first slide was in place) and to report whatever he saw. Each slide was exposed for one minute. When *S* reported movement, he was asked to approximate the extent of movement. Between slides *S* was requested to close his eyes.

Two *S*s never reported movement while the remaining eight *S*s reported movement for every slide. Only the latter *S*s will be considered in the analysis of results. The first slide will be referred to as the adaptation slide.

### 2. *Results*

*a. Latency.* A record was kept by *E* of the latency period, the number of seconds which elapsed before movement was reported for a slide. The adaptation slide yielded latencies of from 1 to 80 seconds, with a mean of 30.5 seconds, the largest mean latency of all slides. Maximum latency occurred for this slide for four *S*s while seven *S*s showed a drop in latency for the next slide.

Excluding the adaptation slide and arranging the remaining five slides in order of increasing size of the perforation, we found an increase in the range of latencies as well as in the mean latency per slide. The ranges were from 2 to 25, 2 to 42, 2 to 45, 3 to 45, and 4 to 90 seconds, with means of 10.9, 12.3, 16.0, 19.7, and 27.8 seconds. There was a +1.00 rank-order correlation between size and latency which, however, did not attain statistical significance at accepted levels of confidence. Comparison of the mean latency of two of the five slides at a time yielded 10 differences, all statistically significant (at or beyond the .05 level); all supported the thesis that latency and size were positively related. If a greater delay before reported movement is considered as an index of a weaker autokinetic effect, then the present findings hint that, with the source of illumination held constant, the larger stimuli tended to yield weaker effects.

*b. Distance.* Of the eight *S*s who reported movement, one refused to estimate the extent of movement. The others displayed considerable individual differences in the distances they reported, which varied from one-tenth of an inch to 12 inches. There were two *S*s who consistently reported movements of less than half of an inch, whereas three others consistently reported movements of two inches or more. Differences between an *S*'s smallest and largest reported extent of movement was less than one inch for some *S*s and over seven inches for others. Of course different *S*s might not have referred to precisely the same standard when they spoke of an inch, or even the same *S* might not have been consistent throughout.

Combining the reports of the seven *S*s who approximated the distance moved, we found that the mean distance was 2.0, 3.6, 3.7, 3.5, 2.6, and 2.3 inches, respectively, for the lights, arranged in order of increasing size. The largest mean distance occurred for the adaptation slide. Of the 15 differences obtained by comparing the mean reported distance for two stimuli at a time, nine differences supported the thesis that the larger the size, the smaller the reported distance. There were five differences which proved statistically significant, of which four favored this thesis. If a smaller reported extent of movement is considered as an indication of a weaker autokinetic effect, then the present findings suggest that, with the source of illumination held constant, there was a slight tendency for larger stimuli to be associated with weaker effects.

## B. EXPERIMENT II

### 1. *Method*

Various modifications were introduced in the apparatus and procedure of the previous experiment.

*a.* Another dark-box was employed, 24 inches wide, 22 inches deep, and 30 inches long, with a four-inch square opening centered on its front panel.

*b.* Ten slides were used. The lengths of the sides of the square perforations were one-sixteenth, one-eighth, one-quarter, one-half, three-quarters, one, one and one-half, two, two and one-half, and three inches.

*c.* A longer darkness adaptation period was allowed. For 30 minutes prior to the start of the session, *S* waited in an anteroom while wearing swimming goggles with red cellophane pasted on the lens.

*d.* In lieu of a shield on *S*'s eyes, all apparatus and furnishings in the experimental room had been covered with black cloth or hidden behind black screens.

*e.* In an effort to make for a more uniform standard of judgment of distance of apparent movement, *S* was given a 12-inch ruler when he entered and, the first time he reported movement for a stimulus, he was asked to approximate the extent of movement by "feeling along the ruler."

*f*. When *S* first reported movement he was asked to mention whenever the light "moves again." This was done to see whether there was any consistent relationship between number of movements and size.

*g*. Each stimulus was exposed for two minutes.

Ten *S*s, including children and adults, were tested individually. For half of them the 10 slides were presented in the order of increasing size of the perforation; then the four-inch square aperture in the front panel was itself illuminated and fixated. For the other *S*s, the 11 lights were presented in the order of decreasing size.

## 2. *Results*

*a*. *Latency; ascending order*. Each of the five *S*s reported movement for some stimuli but none did so for the three largest. If 120 seconds (the two minute exposure period) is taken as the latency period when no movement was reported, then it is found that the mean latencies for the stimuli, in the order of increasing size, were 22.4, 34.8, 27.6, 70.2, 76.3, 72.0, 85.2, 119.8, 120.0, 120.0, and 120.0 seconds, respectively. A rank-order correlation of +.97, significantly different from zero at the .01 level of confidence, prevailed between size and latency. Comparison of the mean latency of two of the 11 stimuli at a time yielded 55 differences, of which 50 favored a positive relationship between size and latency, three were zero, and only two differences were against this relationship. In short, 53 of the 55 differences suggest that, with the source of illumination held constant, latency was a monotonically increasing function of size. Statistical significance was attained by 16 of the differences, all of which supported this relationship.

*b*. *Latency; descending order*. Every *S* reported movements for some lights but none did so for the first two presented. The mean latencies for the lights, in the order of decreasing size, were 120, 120, 85.2, 67.8, 86.0, 91.0, 83.2, 62.4, 51.0, 38.2, and 21.2 seconds respectively. A rank-order correlation of +.89, significantly different from zero at the .01 level, was found between size and latency. Of the 55 inter-stimuli comparisons, 48 supported the thesis that the greater the size, the greater the latency, one was zero, and six were opposed to this thesis. Ten differences, all favoring this thesis, proved statistically significant.

*c*. *Extent of movement; ascending order*. It will be remembered that the first time *S* reported movement for a light, he was asked to estimate the distance by using the 12-inch ruler as a gauge. The mean reported extents of movement for the lights, in the order of increasing size, were .95, .56, .76, .40, .41, .75, .15, .18, 0, 0, and 0 inches. A rank-order correlation of —.90, significant at the .01 level, prevailed between size and mean reported movement. Of the 55 interstimuli comparisons, 46 supported the thesis that greater size was associated with smaller apparent movement, three were zero, and six were opposed to this thesis. Statistical significance was attained by 20 differences, of which 19 favored this thesis.

*d. Extent of movement; descending order.* Combining the reports of the five *S*s to whom the stimuli were presented in the order of decreasing size, we find that the mean reported extents of movement for the lights, in the order presented, were 0, 0, .40, .55, .30, .35, .38, .45, .45 and .51 inches. A rank-order correlation of —.62, not statistically significant at accepted levels of confidence, was obtained between size and reported distance. There were 41 comparisons which favored an inverse relationship between size and reported mean distance and 11 which opposed it while three were zero differences. Of the 18 comparisons which attained statistical significance, 15 supported an inverse relationship between size and mean reported distance.

*e. Frequency of movement; ascending series.* *S* was asked to tell whenever the light "moves again." Combining the responses of the five *S*s who received the lights in the order of increasing size, we find that the total reported numbers of movements were 48, 30, 24, 20, 24, 10, 10, 1, 0, 0, and 0, for the stimuli in the order of presentation. A rank-order correlation of —.98 existed between size and total number of reported movements per slide. An inverse relationship was favored by 49 interslide comparisons and opposed by only one difference while five were zero. That is, 54 of the 55 comparisons favored the hypothesis that, with a constant source of illumination, the number of movements was a monotonically decreasing function of the size of the stimulus. All of the 18 statistically significant differences supported this hypothesis.

If fewer reported movements are considered as indicants of weaker autokinetic effect, then the present findings suggest that, with the source of illumination held constant, larger lights tended to be associated with weaker effects.

*f. Frequency of movement; descending series.* Essentially the same picture was yielded by those who received the stimuli in the order of decreasing size. The total numbers of movements for the 11 lights in the order of presentation, were 0, 0, 8, 9, 9, 8, 12, 21, 48, 57, and 95, respectively. A rank-order correlation of —.95, significant at the .01 level, was obtained.

There were 50 interstimuli differences which favored an inverse relationship, and only two were contrary to the trend while three were zero. The five statistically significant differences supported an inverse relationship.

*g. In summary.* Whether the lights were presented in the order of increasing size or of decreasing size, the data relative to latency, extent, and frequency of reported movement suggested that the larger stimuli occasioned weaker autokinetic effects.

## C. Experiment III

### 1. *Method*

Various modifications were introduced in the apparatus and procedure of Experiment II.

*a.* The opening in the box's front panel was a five-inch square rather than a four-inch one.

*b.* Instead of one 100-watt bulb, two 60-watt bulbs inside the box served as the source of illumination. They were in the two lower front corners of the box.

*c.* The slides were those of Experiment II but instead of being presented in consistent order, with reference to size of the perforation, they were shuffled into a random order and presented in this same order to all *S*s. The lengths of the sides of the square perforations, in the order of presentation, were one-quarter, one and one-half, one-eighth, two, one-sixteenth, three-quarters, one-half, one, and two and one-half inches. The eleventh stimulus was the five-inch square aperture in the box's front panel.

*d.* There were two *E*s, one of whom presented the slides while the other handled the stopwatch and the data sheets.

*e.* The first time *S* mentioned movement for the first slide presented, the adaptation slide, he was told that henceforth he was to report continuously on the kind, direction, and extent of movement, and that he was always to use the ruler in gauging the distance moved.

Twenty *S*s, aged from 10 to 16 years, were tested individually.

### 2. *Results*

*a. Kinds of movement.* Two different kinds of movements were noted: movements toward *S* or away from him as well as movements in a plane perpendicular to *S*'s line of sight. To distinguish between the two, we shall refer to the former as forward-backward or *F-B* movement and to the latter as autokinetic or *A*-movement. Latency periods as well as other data were determined separately for the two kinds of movements.

*b. Latency.* *F-B* movements tended to be seen earlier than *A*-movements. For every stimulus the mean latency for *F-B* movement was less than that for *A*-movement. Usually once *A*-movements occurred, *F-B* movements ceased but in some cases, particularly for the larger lights, *F-B* reoccurred after *A*-movement ceased.

We were particularly interested in the direction of the first movement announced for a light. The 20 *S*s had an opportunity to make a total of 220 such 'first" announcements for the 11 stimuli but there were 20 instances of no reports, leaving only 200 such announcements. Of these, 86 reported *F-B* movement and 114 *A*-movement. Further subdivision shows that movements towards *S* were mentioned in 47 cases, away from *S* in 39 cases, movements to the left in 26 cases, to the right in 23 cases, upwards in 48, downwards in 12, and both to the side and up or down in only 5 cases. Excluding the adaptation slide, we find that for the five smallest lights the first reports of direction included 30 *F-B* and 65 *A*-movements while for the five largest they included 45 *F-B* and 41 *A*-movements.

The relationship between latency and size has been analyzed only for *A*-movement, with 120 seconds serving as the latency period if such movement was not reported for a particular stimulus. If we consider the lights in the order of increasing size (which was not the order of presentation), then we find that the mean latencies were 22.0, 24.8, 61.5, 25.8, 42.1, 35.0, 54.0, 35.3, 35.5, 54.9, and 66.7 seconds. The highest latency was for the largest light (the five-inch square aperture) and the next largest for the adaptation slide. Excluding the adaptation slide, we obtain a rank-order correlation of +.87, approaching significance at the .01 level. This corroborates the positive relationship between latency and size revealed by Experiments I and II. Comparison of two lights at a time (with the adaptation slide excluded) yielded 45 differences of which 40 supported a positive relationship.

*c. Extent of movement.* While there was no consistent relationship between size and reported distance, there was a slight, statistically insignificant, tendency for the larger lights to be reported as moving smaller distances. Thus the mean distance noted for the five smallest (adaptation slide excluded) was 1.5 inches as compared to 1.2 inches for the five largest.

*d. Frequency of movement.* Data pertaining to the number of times each light was said to move were calculated separately for *A*- and for *F-B* movement. For every slide, but not for the box's five-inch square aperture, *A*-movements were more often reported than *F-B* movements. All together, the 20 *S*s reported 2,215 *A*-movements and 551 *F-B* movements for the 11 stimuli. Considering the lights in the order of increasing size, we find that the numbers of *S*s who announced *A*-movements more frequently were 20, 18, 16, 18, 15, 16, 13, 12, 13, 9, and 9, respectively, while the numbers of *S* who reported both kinds of movements the same number of times or who failed to report either were 0, 2, 2, 2, 3, 2, 2, 3, 4, 2, and 7 respectively. This means that only a few *S*s announced *F-B* movements more frequently and that they did so mainly for the larger stimuli.

Analysis of the relationship between size and frequency of movement was undertaken only for *A*-movements. The total numbers of such movements reported by the 20 *S*s for the lights, considered in the order of increasing size, were 420, 305, 244, 268, 216, 210, 120, 146, 149, 101, and 36. As in the previous experiments, an inverse relationship prevailed between size and frequency of movement. A statistically significant correlation of —.95 was obtained. Of the 55 comparisons of two lights at a time, all of which proved significant (most of them at the .01 level), 51 supported an inverse relationship.

## D. Experiment IV

### 1. *Method*

The procedure of Experiment III was used with 10 new *S*s, aged from 10 to 14, and again both *A*- and *F-B* movements were reported. Trends of results again confirmed the hypothesis that, with the source of illumination constant, larger stimuli tended to give rise to weaker autokinetic effects.

### 2. *Results*

*a. Latency.* For every light the mean latency for *F-B* movement was less than that for *A*-movement. The mean latencies for *A*-movements for the 11 lights considered in the order of increasing size (which was not the order of presentation), were 18.2, 16.3, 58.9, 19.8, 16.2, 20.0, 46.0, 21.7, 41.5, 46.2, and 85.1 seconds. Exclusion of the adaptation slide (latency of 59.9 seconds) yielded a rank-order correlation of +.81. That greater latency was associated with greater size was attested to by 36 of the 45 comparisons.

*b. Extent of movement.* The mean reported distances of *A*-movements for the stimuli, in order of increasing size, were 9.0, 8.3, 7.5, 4.1, 2.2, 1.6, 2.8, 3.9, 5.1, 1.0, and 1.1 inches. While there was no consistent relationship between size and distance, there was a slight tendency for larger lights to be reported as moving smaller distances. For the five smallest lights the mean reported distance was 6.2 inches while for the other lights it was 2.6 inches.

*c. Frequency of movement.* Again an inverse relationship tended to prevail between size and frequency of reported *A*-movements. The total number of *A*-movements announced by the 10 *S*s for the lights, considered in the order of increasing size, were 154, 118, 770, 101, 79, 97, 45, 59, 53, 42, and 21, respectively. Excluding the adaptation slide (770 movements), we obtained a rank-order correlation of —.95, differing significantly from zero at the .01 level. An inverse relationship was supported by 42 of the 45 interstimuli comparisons, with the adaptation slide omitted, and by 49 of the 55 comparisons when this slide was included. Of the 53 statistically significant differences, 47 favored an inverse relationship.

## E. Experiment V

### 1. *Method*

For most *S*s of Experiments III and IV, the session did not end with the presentation of the 11 lights. After the eleventh, various modifications were introduced in attempts to study the relationship between size of the stimulus and (*a*) rate or perceived movement; (*b*) suggestibility; (*c*) voluntary cessation of movement, and (*d*) variations in illumination. The modifications will be described in conjunction with the findings they yielded.

### 2. *Results*

*a. Speed of movement.* After the eleventh stimulus, *S*s were questioned about the rate at which the lights had moved. Most *S*s stated that the smaller lights had tended to move faster; this confirmed their previous comments during the session. As a further check on the relative rates of movements, the smallest stimulus (1/16 of an inch square perforation) and the largest (5-inch square aperture) were again presented to the 10 *S*s of Experiment IV. Every *S* commented on the greater speed of the smaller light. The reactions can be summed up in the comment, "The big one is lazy, sluggish, but the small one jumps around." If faster apparent movement is taken as a token of a stronger autokinetic effect, then this finding substantiates the thesis that larger stimuli tended to occasion weaker effects.

*b. The rôle of suggestion.* After the eleventh stimulus of the main procedure, 18 *S*s of Experiment III were shown both the smallest and the largest lights. For each of these *E* made the following suggestions:

> I'm going to move this light up. Ready now. There! Can you see it? Up, up, up! See it? . . . Now I'm going to move it like the bottom of a ball. Left, right, left, right. See it? . . . Now I'm going to make it bounce up and down like a yo-yo. Up, down, up, down. See it? Now the light will move to form a giant figure 8. There it goes. See it?

If *S* did not accept a suggestion, the other *E* added, "Can't you see it? It is going up and down (or to right or left, etc.). Look carefully. Pay attention."

For half of the *S*s, the procedure was first carried out with the smallest light and then with the largest while for the others the reverse order was used. The order seemed to have little influence on the results. Individual differences in apparent suggestibility were observed, with some *S*s being readily induced to see the suggested movements while others would not at all comply with the suggestions. But, in general, it was easier to obtain agreement or at least partial agreement with the proferred description for the smaller stimulus. Of the 18 *S*s, 10 proved rather amenable to suggestion for the small light but only four for the larger one. A similar trend was found when this procedure was employed with the 10 *S*s of Experiment IV (after the variation pertinent to rate of movement).

The smaller light's greater susceptibility to the suggested pattern of movement might be interpreted as evidence of a stronger autokinetic effect, one with a more powerful deceptive influence on *S*. If this interpretation is allowed, then the present findings fit the trend that smaller stimuli and stronger effects tend to go hand in hand, given a constant source of illumination.

*c. Voluntary cessation of movement.* We wondered whether voluntary cessation of the autokinetic effect could be achieved by *S* and whether the size of the light would prove to be a differential factor. Therefore, after the phase on suggestibility, another modification was introduced for eight *S*s of Experiment III. For the slides with the 1/16 of an inch square perforation and with the 3-inch square perforation, *S* was asked, after he reported movement, to "try to stop the movement by doing something with your eyes but you must not move your head nor are you allowed to close your eyes." For half of the *S*s, the smallest light was the first to be presented and for the others the larger was first. The order of presentation did not seem to be an important determinant of results.

Two Ss said that they could not stop the movements of either light. Two other Ss claimed that they succeeded in stopping the movements of both lights; one said he did so by "staring hard" and another said he succeeded by casting his eyes downward while keeping his head immobile. The remaining four Ss contended that they could stop the movements of the larger light but not of the small one.

This finding also might be interpreted as evidence of a stronger autokinetic effect for the smaller stimulus, one from which S could not so readily break away under the described conditions.

*d. Variations in illumination.* For the 10 Ss of Experiment III a new phase was conducted immediately after the study of suggestibility. The illumination of the surface in which the *A*-light was embedded was gradually altered. Apparatus employed for this purpose has been described in another report (1). Instead of placing the slide on the front panel of the box, the top of the box was opened and the slide was placed vertically so that it divided the box in half. In the rear compartment a 40-watt bulb was on continuously and served as the source of illumination for the perforation in the slide. The two 60-watt bulbs in the front compartment were connected to a variac which regulated the amount of incoming current. The variac, General Radio Company's V-5, had a dial which was calibrated from 0 to 135 volts. A zero dial setting represented no illumination furnished by the two 60-watt bulbs and a 135 dial setting represented the maximum illumination they could furnish. By turning the variac dial, which was on *E*'s table, *E* was able to change the illumination of the box, and hence, of the slide containing the *A*-light.

For half of the Ss the procedure to be described was applied first to the slide with the 1/16-inch square perforation and then to the slide with the 3-inch square perforation while for the other Ss the order of the two slides was reversed. When the slide was first presented, the room was in darkness except for the perforation in the slide which was illuminated by the 40-watt bulb behind it. The *A*-light was visible to *S* through the opening in the box's front panel. One minute after *S* reported movement, *E* began gradually to increase the dial setting, thereby increasing the illumination of the box's interior. *S* was of course not informed that the illumination was being changed. We were interested in the dial setting at which *S* would first report cessation of movement (the cease-setting) and in whether the size of the *A*-light would be a differential determinant.

One *S* reported cessation of movement for the two *A*-lights at about the same dial setting. For the nine other *S*s the cease-setting was lower for the larger *A*-light; that is, the larger stimulus required less illumination to be furnished by the two 60-watt bulbs. All 10 *S*s reported cessation of autokinetic movement for the larger stimulus but three *S*s continued to report movement for the smaller one throughout the whole range of the variac. Even when the variac dial was set at 135, and the entire room was well illuminated by the light streaming from the box's front aperture, these three *S*s claimed that they saw the small *A*-light moving. The 10 *S*s' cease-settings for the large stimulus ranged from 0 to 96, with a mean of 49. If 135 is considered as the cease-setting for those who continued to report movement, then it is found that the cease-settings for the small stimulus ranged from 40 to 135, with a mean of 86. The illumination furnished by the two 60-watt bulbs at the cease-setting was about 40 foot-candles less, on the average, for the larger stimulus than for the smaller one. Unfortunately, the total illuminosity was not measured so that the differential contribution of the two stimuli were not taken into account. With the 40-watt bulb as the source, the 3-inch square perforation of course made a greater contribution to the total illuminosity than did the 1/16-inch square perforation.

If a lower variac cease-setting is taken as evidence of a weaker autokinetic effect, then the present findings suggest that the larger stimulus was subject to the weaker effect. In this respect the results fit in well with the trends established by all other indices which we have discussed.

## F. Summary and Discussion

A number of exploratory experiments, administered individually to 50 *S*s, studied the relationship between autokinetic effect and the size of the stationary light in a dark room (the size of the *A*-light or stimulus). In one experiment six stimuli were used, ranging in size from a square each of whose sides was one-eighth of an inch in length to one whose sides were each one inch long. In other experiments, 11 stimuli were used, ranging in size from a square each of whose sides was one-sixteenth of an inch in length to one whose sides were each five inches long. Illumination was furnished by one bulb or by two bulbs in the dark box which were on continuously throughout the session so that, within each experiment, there was a constant source of illumination.

The stimuli were presented in the order of increasing size or in the order of decreasing size or in random error. Regardless of the order of presentation, certain trends prevailed. A positive relationship between size and latency (the time which elapsed before movement was reported) was attested to by rank-order correlations ranging from +.81 to +1.00 yielded by four experiments (usually with the first light, the adaptation light, excluded); most of these correlations were significant at or near the .01 level of confidence. Moreover, when the latencies of two lights at a time were compared, from 80 to 100 per cent of the resulting comparisons favored the thesis that the larger the light, the greater the latency. An inverse relationship between size and frequency of reported movements was supported by rank-order correlations of from —.95 to —.98, all differing significantly from zero at the .01 level of confidence. From 90 to 93 per cent of the interstimuli comparisons were in line with the thesis that the greater the size, the less frequently movements tend to be reported. There was also some tendency in each of the four experiments, for larger lights to be associated with smaller extents of reported movements.

These data were interpreted as implying that, with the source of illumination held constant, the larger the stimulus, the weaker the autokinetic effects tended to be. Further evidence along the same lines was obtained when some *S*s participated in a number of modifications, subsequent to the main procedure, in which only two lights were again presented, the smallest and one of the largest. The smaller light was found (*a*) to be reported as moving at a faster rate of speed, (*b*) to be more susceptible to directions and patterns of movement suggested by *E*, (*c*) to be less susceptible to voluntary cessation of movement on *S*'s part, and (*d*) to require more illumination to be furnished by supplementary bulbs in the box in order for the *A*-light to appear stabilized.

When *S* was asked to report continuously on what he observed, two kinds of movements were noted: movement toward or away from *S* (called forward-backward or *F-B* movement) and movement in a plane perpendicular to *S*'s line of sight (referred to as autokinetic or *A*-movement). *F-B* movements tended to occur earlier than *A*-movements and to be more prevalent for the larger lights than for the smaller ones. Analysis of *A*-movements showed a tendency for a direct relationship between size and latency and an inverse relationship between size and frequency as well as between size and extent of movement.

Since the source of illumination (e.g., a 100-watt bulb within the box) remained constant within each experiment, and since the area through which the light could stream varied with the size of the stimulus, we cannot be certain whether the present findings are demonstrating the influence on the autokinetic effect of the size of the stimulus or of the total illumination or of both. As a preliminary attempt to untangle the factors of size and of total illumination, one might vary the source of illumination (e.g., by attaching it to a rheostat) so that as the stimuli became larger, the source of illumination would become weaker, to the extent that the total illumination would always remain the same; but this procedure has the deficiency that the intensity of illumination per unit of area is not constant. In addition to further experiments which would reveal the separate contribution of the size of the *A*-light, the intensity of illumination per unit of area, and the total illuminosity, there is need for investigations of the influence of the visual angle, of the shape of the *A*-light, and of the influence of the perimeter of the stimulus as compared to the area.

In brief, all indices which we studied suggested that as the size of the square perforations varied, with the source of illumination held constant, there was a very decided tendency for larger perforations to occasion weaker autokinetic effects.

## REFERENCE

1. LUCHINS, A. S. The relation of the autokinetic effect to the illumination of the surrounding field. *J. Gen. Psychol.*, 1954, **50**, 29-37.

*Department of Psychology*
*McGill University*
*Montreal, Canada*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 453-456.

# SOME CIRCUMSTANCES SURROUNDING APPARENT MOVEMENT IN THE LINE OF REGARD*

*Department of Psychology, Michigan State College*

---

S. HOWARD BARTLEY AND JAMES W. MILLER

---

## A. PROBLEM

The study of visually perceived motion in the line of regard has been somewhat neglected. Third-dimensional apparent movement may be produced by (*a*) the successive presentation of stationary targets at different *distances;* (*b*) the successive presentation of targets of different *size,* but at the same distance; and (*c*) by manipulating the intensity of a stationary target (5).

The present paper reports the investigation of apparent movement in the line of regard produced by method (*b*). We decided to call movement in the line of regard *adab* movement, to give it a short label and distinguish it from all other movement possessing a third-dimensional component. *Adab* is a combination of the two prefixes *ad* and *ab,* having to do with toward and away from.

## B. APPARATUS AND PROCEDURE

A Harvard tachistoscope was used to provide virtually instantaneous shifts from one target to the other, and exposures of the second target for controlled lengths of time. The targets were lighted squares presented in a dark field, provided by openings in partitions placed across the alleys in the instrument, and passing light from the opal glass surfaces beyond.

The experimental procedure consisted in presenting one target and then momentarily interrupting it by substituting the second target for any one of the various lengths of time. A complete run consisted in ascertaining the whole gamut of significant changes that was produced by manipulating second-target duration. For final use, only those phenomena that seemed to have a relatively definite emergence point were selected for consideration.

Binocular viewing was used throughout. The distance of the targets from the eye was 41 cm. The size of the first, or Target *N,* was 2.6 cm. When size was manipulated, the other squares were 1.3; 1.9 and 2.27

---

*Received in the Editorial Office on July 1, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

cm., *A, B,* and *C* respectively. When the second square was the same size as the first, it was known as Target *D*. The duration of the second target ranged from 10 msec. to 700 or 1000 msec., in gross steps. Three intensities of the second target were used; namely, 94 per cent, 16 per cent, and 3 per cent of that of the first target.

## C. RESULTS

Adab movement seems to depend upon several factors that were manipulated in the investigation. One of these is the intensity of the second target in relation to the first. In all but one of the 14 conditions in which adab was elicited, for the two observers, the weaker the second target the longer its exposure had to be. These exposures ranged from 75 msec. to 300 msec. for Observer *B,* and 100 to 250 msec. for Observer *M*.

A second factor was that of second-target size with respect to that of the first. There was a tendency for larger targets to require longer exposures to produce adab movement. The rôle played by size of second target was not as clear as the factor of intensity, possibly owing to the increase in the size of the second target involving two things: (*a*) a greater amount of light due to a greater area, and (*b*) the consequent bringing of the image of the second target border closer to retinal areas just occupied by the first target. Whereas (*a*) would tend to elicit a quicker emergence of square two, (*b*) would operate in the opposite direction as suggested by the earlier work of Fry and Bartley (4), and also of Werner (6). They found that border processes may work against each other.

For Observer *B*, for Target *A,* no adab movement was ever elicited. For Observer *M,* it was elicited when *A* was at the two lesser levels of intensity. With Target *B,* adab movement was elicited at about the same target durations for both observers; at 100 msec. for the more intense level, and at 175 msec. for the lowest level.

With Target *C,* which was approaching the size of *N,* the range ran from 75 msec. to 300 msec. for Observer *B,* and from 100 msec. to 250 msec. for Observer *M*.

A circumstance surrounding the emergences of adab movement seemed to be the instability of the border of *N,* as represented in the gamma movement involved as *N* was re-established after the termination of the second target.

The behavior of *N* when interrupted and not followed by another square was taken as the basis for comparing the behavior of *N* when followed by various second targets. When *N* was interrupted, gamma movement was produced, but the radial centripetal and centrifugal movements tended to occur within the bounds of *N* itself, and not to involve the borders until the duration of the interruption became fairly long. When the interval reached about 380 msec. for Observer *B*, and about 410 msec. for Observer *M*, the centrifugal movement, as *N* reappeared, involved a shift of *N*'s contour. This we shall refer to as instability of the border or contour, for up to this point the border merely disappeared and reappeared without any sidewise or radial movement whatsoever.

When *N* was replaced by a second target, be it *A*, *B*, or *C*, some illumination was put into the field which otherwise would be unlighted during the interval of *N*'s absence. It was thought that this new target illumination might partially substitute for the light that was withdrawn when *N* was interrupted, helping sustain the contours of *N*. Just the opposite happened. With the second target following *N*, it took a much shorter interruption of *N* to produce instability. In fact the more intense the second target, the sooner the border of *N* became unstable. The times for the instability vary from 125 to 150 msec., for Target *A* (Observer *B*). For Target *B*, the range ran from 100 to 150 msec., and for Target *C*, from 100 to 225 msec.

Apparently the potent factor in *N*'s contour instability was the putting of a new border into the area from which *N* was removed (2, 3, 4). The process that produced the new border depended for its effectiveness upon the intensity of the second target and its size. The larger the second target, the nearer the new border was to the old borders of *N* which had just disappeared.

In nearly every case, the initial contours of the second target were set up before the border of *N* became very noticeably unstable. For Observer *B*, this interval ranged from 115 to 25 msec. for Target *A*; from 90 to 50 msec. for Target *B*; and from 90 to 0 msec. for Target *C*. This suggests that the nearer the contour of the second target was to the contour of *N*, the shorter the interval till the contours of *N* became unstable. The opposite relation held for the *intensity* of the second target. The less intense, the shorter the interval between contour formation of the second target and the instability of *N*. Rather than being truly paradoxical, this was an incidental relation dependent upon the fact that with reduced intensity of the second target the longer it took for its contours to emerge.

The findings show that: (*a*) The initial contours of the second target square emerge earlier than other events in the sequence observed. (*b*) Contour emerges later for targets of lesser intensity. (*c*) The initial contour of the second square is accompanied by a tapered dark area just outside of it which gradually fades and disappears with the further lengthening of the duration of the second target (1). (*d*) Instability of *N*'s contour emerges generally much later than the emergence of the contour of the second square. (*e*) The margin of time between the two lessens as the intensity of the second target is reduced. The contour of the second square is very much slower in emerging with weak second targets. (*f*) Adab movement does not occur under all conditions and not equally marked when it does occur. Adab movement does not occur when the conditions for discrepancy between the emergence of contours of Square 2 and the instability of the contour of *N* is great. This discrepancy is maximum for Target *A* at its maximum intensity. It is here that neither observer saw adab movement. As the intensity of Target *A* is reduced, the discrepancy became less and Observer *M* began to see adab movement. Observer *B* did not see adab movement with Target *A*. (*g*) When adab movement appeared, it first emerged with durations of Target 2 adjusted about to produce instability of the first square. The discrepancy when it occurs was never more than 25 msec. That is, the duration of Target 2 was never 25 msec. shorter than for the instability of Target 1. It could be longer. In fact with the weakest intensity of Target *C*, adab did not occur until a value of 75 msec. was obtained. Although our conditions did not test it, the expectation is that if the second target is made very weak, adab movement would never occur at all. The target would never emerge as a contoured square no matter how long the duration of presentation.

## REFERENCES

1. BARTLEY, S. H. Brightness enhancement in relation to target intensity. *J. of Psychol.*, 1951, **32**, 57-62.
2. FRY, G. A. The relation of border contrast to distinctness of vision. *Psychol. Rev.*, 1931, **38**, 542-549.
3. ———. Depression of the activity aroused by a flash of light by applying a second immediately afterwards to adjacent areas of the retina. *Am. J. Physiol.*, 1934, **108**, 701-707.
4. FRY, G. A., & BARTLEY, S. H. The effect of one border in the visual field upon the threshold of another. *Am. J. Physiol.*, 1935, **112**, 414-421.
5. NEWMAN, E. B. Versuche über das Gamma-Phenomen. *Psychol. Forsch.*, 1934, **19**, 102-121.
6. WERNER, H. Studies in contour. *Am. J. Psychol.*, 1935, **47**, 40-64; 1940, **53**, 418-422.

*Department of Psychology*
*Michigan State College*
*East Lansing, Michigan*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 457-466.

# PERFORMANCE ON THE SAME PSYCHOMOTOR TASK UNDER DIFFERENT STRESSFUL CONDITIONS*[1]

*Worcester State Hospital and the Worcester Foundation for Experimental Biology*

OSCAR A. PARSONS, LESLIE PHILLIPS, AND JOHN E. LANE[2]

## A. THE PROBLEM

The prediction of performance under difficult or trying conditions is an important problem area in psychology. In an effort to attain greater understanding of behavior under such conditions, many investigators have used concepts such as "frustration tolerance" (15), "intellectual control" (1, 18), "stress tolerance" (17) and the like. An assumption which appears to be implicit in the use of such conceptual variables is that individuals can be ordered on a scale of resistance to difficult or stressful conditions. It is often suggested that individuals who score high on such variables and, therefore, on this hypothesized scale will perform more adequately in a variety of stressful situations.

However, most investigators in the area of stress experimentation have been content to examine performance under one type of stressful condition. The empirical question arises: to what extent can one generalize from performance in any one experimental stress situation? This in turn emphasizes the basic theoretical question: is there a general factor of resistance to stressful conditions?

Inferential support for the presence of a general factor can be educed from clinical and experimental studies in psychopathology. Schizophrenics, for example, perform more poorly than normal *S*s under many different experimental conditions (9). Since the schizophrenic is usually thought to be a person who fails to meet the pressure of everyday life, even in his premorbid behavior, one might infer a general inability to cope with stressful conditions (14). Malmo *et al.* have demonstrated the tendency of neurotics, as com-

---

*Received in the Editorial Office on July 2, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a research grant from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

[2] We are indebted to Dr. Fred Elmadjian of the Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts, for the supervision of the eosinophil determinations and his aid in the interpretation of this measure. Thanks are also due Dr. Austin W. Berkeley for the design and modification of much of the equipment used in this experiment.

pared with normals, to manifest motor disturbances under several stressful conditions (11, 12).

On the other hand, studies with normal *S*s have given equivocal results. In some experiments performance under frustrative or stressful conditions has been found to be related to more general variables such as "intellectual control" or "adjustment." In other experiments under similar conditions, such relationships have not been confirmed. Lazarus, Deese, and Osler (10), and Deese and Lazarus (2) summarize these experiments.

Undoubtedly many factors might play a rôle in discrepant results, e.g., differences in populations, surroundings, performance skills, intelligence, etc. These variables might serve to obscure any general factor of resistance to different stressful conditions. A more adequate test for such a general factor would be to examine the relationship between performances of a relatively homogeneous group who perform upon the *same* task under objectively *different* stressful conditions. The present experiment was designed to make such a test.

One additional variable would appear to be important. Lazarus, Deese, and Osler (10) have suggested that the usual assumption of equal involvement on the part of *S*s under different experimental stress conditions is seldom warranted. Thus, if *S*s were involved in one type of stressful situation but were relatively uninvolved under a different set of stressful conditions, performances in the two situations might be unrelated. It would seem desirable, therefore, that some relatively independent and objective criterion of involvement be utilized. An experiment by Stopol, reported in abstract form, is relevant (17). This investigator examined performances upon the same task, the digit symbol test, under "social" and "mechanical" stress conditions. No relationship between performances was found; however, no control was reported for the involvement of *S*s in the two situations.

Attempts to evaluate involvement in stress situations have typically employed impressionistic techniques, e.g., ratings of overt behavior. To attain a less subjective and more easily quantifiable measure of involvement, a physiological variable, the level of circulating eosinophils in the blood, was chosen. While the physiological mechanisms involved are not clearly understood, numerous studies have shown the responsiveness of this measure to stressful situations. Thus Selye (16) reports that lowered eosinophil levels have been demonstrated in students following examinations, in usually anxious children who were afraid of an approaching injection, in telephone operators working under trying conditions, and in *S*s who had to solve difficult mathematical problems. It is obvious that physiological responsivity under such conditions would not occur unless the *S*s were involved psychologically, since the actual physical demands in these situations were minimal.

Thus two questions are asked in this experiment: (*a*) will performance under one type of stressful experimental conditions be positively related to performance on the same task under an objectively different type of stressful conditions? (*b*) Will these results be affected by the involvement of *S*s as measured by the eosinophil responsivity scores?

## B. METHOD

### 1. *Population*

The *S*s were 57 white, male, industrial workers, most of whom were rated as skilled or semi-skilled, and employed by the same company. Almost all were married, had children, and were maintaining and providing for their families. No gross psychopathological features were observed and none had a history of hospitalization for mental disease. Table 1 lists the pertinent information concerning chronological age, mental age, and educational level.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF CHRONOLOGICAL AGE, MENTAL AGE,* AND EDUCATIONAL LEVEL OF *S*s ($N = 57$)

| | | *Mean* | *SD* |
|---|---|---|---|
| Chronological Age | (Years) | 39.0 | 4.09 |
| Mental Age | (Years) | 14.8 | 2.80 |
| Educational Level | (Years completed) | 10.1 | 2.32 |

*Derived from Stanford-Binet Vocabulary (1916).

### 2. *Performance Task*

A modified version of the Dunlap hand-steadiness apparatus (4) was used in this experiment. The specific task was to keep a stylus from touching the sides of a hole one-eighth inch in diameter. The stylus was fixed so that it could not be withdrawn from the hole and at the same time was so balanced that it required *S*'s constant attention to prevent it from touching the sides. Each time the stylus touched the sides, an electrical contact was made which shut off a small light that was otherwise illuminated. In this way *S* knew when he was committing an error. An electric timer provided a measure of error-free time for each trial, i.e., that time in which the stylus was not contacting the sides of the hole.

### 3. *Experimental Conditions*

On the basis of the experimental conditions Lazarus *et al.* distinguish between two main types of stress experimentation: task-induced stress and failure-induced stress (10). Task-induced stress includes those experiments in which the task itself makes excessive demands upon *S*, e.g., rapid pacing or introduction of distracting stimuli during the course of the performance. Failure-induced stress refers to those experiments wherein the experimenter attempts to have *S* experience failure through the introduction of false norms, caustic comments, etc. The two different experimental conditions utilized in this experiment can be subsumed under the categories distinguished above and will be labeled: "distraction stress" and "failure stress", stress being defined in terms of the stimulus conditions.

## 4. *Distraction Stress Conditions*

*S* was informed that he was to keep the pointer away from the sides of the hole. He was then told that certain things would happen to make his job more difficult, and that he should try to keep the light on no matter what might happen.

After a practice trial, *S* was given 30 trials of eight seconds duration each, with a half-minute rest between trials. During 10 of the last 20 trials distracting stimuli, consisting of airblasts, bright flashing overhead light, loud horn and bell, were introduced according to a pre-arranged irregular pattern that was held constant for all *S*s.

Four measures of performance under these conditions were derived.

*a. Prestress level.* The mean error-free time of the last three trials which preceded the introduction of distracting stimuli. (The last three prestress trials represented *S*'s functioning level at the time stress was introduced, hence served as the base line for comparison with the stress trials.)

*b. Overall stress score.* The mean error-free time for the 20 trials following the introduction of distracting stimuli, corrected for the prestress level.[3]

*c. Distraction stress score.* The mean error-free time for the 10 trials during which distracting stimuli were introduced, corrected for prestress level.

*d. Distraction free score.* The mean error-free time for those 10 trials which were interspersed among distraction trials but during which distracting stimuli did *not* occur, corrected for prestress level.

---

[3] Two problems are usually encountered in trying to achieve a stress score for the individual: (*a*) correlation with prestress level, and (*b*) regression to the mean tendencies (10). All stress measures used in this experiment were corrected for the possible confounding effect of these two factors by means of the regression equation.

## 5. *Failure Stress Conditions*

*S*s were similarly instructed to keep the pointer from touching the sides of the hole. In contrast to the distraction stress situation, *S*s were told that they were not performing as adequately as others of their own occupational level. Each *S* was told that he would receive an "improvement score" which would then be compared with one hundred other people of his own occupational status.

Actually, scores reported to *S* followed a prearranged pattern which was approximately the same for all *S*s. To impede realistic evaluation of performance, scores were reported after every two trials. *S* was given 16 performance scores representing a total of 32 eight-second trials. To maximize involvement, aspirations were obtained for each trial. The first seven trials were designed to be a success phase. *S* was told that his improvement score was above average (the 50 per cent mark) on each trial. The general trend of the reported scores was towards decidedly better performance. Two transitional trials followed and then on each of the last seven trials, *S* was informed that his scores were below the average. During this period the general trend of the reported scores was towards a definite decrement in performance. Thus, on the failure trials each *S* not only scored below all his success trials but also below the group with whom he was being compared. During the test the reported performance was plotted on a graph placed in front of *S*. All scores above the 50 per cent line were plotted in black and all scores below were plotted in red. *S*'s score was reported to him in approximately the following form: "That time you beat .................... people out of a hundred, as far as your improvement score is concerned. How many people do you think you will beat next time?" Aspirations and performance were recorded after each trial.

Two measures of performance under these conditions were derived: (*a*) Prestress level: the mean error-free time on the last three success trials. (*b*) Failure score: the mean error-free time for the trials under failure conditions corrected for the last three success trials.

## 6. *Eosinophil Measure*

Intravenous blood samples were taken 15 minutes before and 15 minutes after the stress situations. The circulating eosinophil levels were determined by a modification of the Dunger method in accordance with suggestions by Discombe (3) and Forsham, Thorn, Prunty, and Hills (8).

Raw values were plotted and found to be skewed. Transformation into log values gave rise to approximately normal distributions. A plot of poststress against prestress log values revealed uniform variability along the arrays with very little regression to the mean effect. A responsivity score for each individual was obtained by subtracting his prestress log value from his poststress log value. Thus negatively signed scores represent greater drop in eosinophil level and are in the direction of greater physiological responsivity.

### 7. *Experimental Design*

The overall design was such that approximately half of the *S*s had the distraction stress situation on the first day and failure stress situation on the second day. This was labeled the "Distract-Fail" group ($N = 27$); the order was reversed for the other *S*s, who were called the "Fail-Distract" group ($N = 30$).

## C. RESULTS

A comparison of changes in circulating eosinophil levels as a function of the two stressful conditions is presented in Table 2. Under both conditions, a statistically significant drop in stress level is found. The mean drops are practically identical. From this it would seem reasonable to infer that the group as a whole was involved in both situations.

TABLE 2
COMPARISON OF CHANGES IN LOG VALUES OF EOSINOPHIL LEVELS IN THE TWO STRESS SITUATIONS

|  | $N$* | $\bar{x}$ difference** | S.E. of $\bar{x}$ | 't' test for correlated means | $p$ |
|---|---|---|---|---|---|
| Failure stress | 49 | —6.06 | 1.25 | 4.84 | $<.001$ |
| Distraction stress | 49 | —6.00 | 1.54 | 3.90 | $<.001$ |

*Only subjects who had reliable eosinophil counts on both stresses were used.
**The difference score was calculated by subtracting prestress level log values from stress level log values and multiplying by a constant (one hundred).

Pearson product-moment correlations were calculated for the pertinent relationships between the performance scores from the distraction stress situation and the failure stress situation. The results are presented in Table 3. Consideration should be given first to the relationship between prestress levels of performance, under the two different conditions. If these prestress levels were not significantly related, interpretation of stress relationships, or lack of relationships would be less clear. Thus it is important to note that, despite extensive differences in psychological conditions under which *S* was performing, a significant positive correlation ($r = +.33$, $p = .01$) between prestress levels obtains. However, when stress scores, corrected for prestress level, are correlated, no significant relationship is found. Indeed, all the correlations are close to zero.

### TABLE 3
### PRODUCT MOMENT CORRELATIONS BETWEEN PERFORMANCE SCORES UNDER FAILURE STRESS AND DISTRACTION-STRESS CONDITIONS ($N = 57$)

| | Pearson $r$ | $p$ |
|---|---|---|
| Failure prestress level vs. Distraction prestress level | .33 | =.01* |
| Failure stress score vs. Distraction overall stress score | .00 | — |
| Failure stress score vs. Distraction stress score | —.10 | — |
| Failure stress score vs. Distraction free score | —.02 | — |

*Evaluated by Fisher's $r$ to $z$ transformation (5).

Before interpreting these results several points must be discussed. We have noted that the group responded with approximately the same eosinophil drop in the two stress situations. The inference was made that the group, as a whole, was involved in both situations. It may be, however, that *individual* Ss were not consistent in their response to the two situations. Ss who were responsive in one stress situation might have been relatively unresponsive in the other. Accordingly, a correlation between the eosinophil responsivity scores was calculated. The result ($r = +.20$, $p > .05$) indicated that Ss were not consistent in their eosinophil response to the two stress situations. The possibility existed, therefore, that different performance relationships might obtain in sub-groups of Ss, classified on the basis of their eosinophil response. To examine this the total sample was divided into three groups: (*a*) Consistent responders—those Ss who were above the median eosinophil responsivity score for each stress situation ($N = 14$). (*b*) Consistent non-responders—those who were below the median eosinophil responsivity score for each situation ($N = 15$). (*c*) Inconsistent responders—those who were above the median in one stress situation and below the median in the other ($N = 20$).

Spearman rank-order correlations (Rho) were then calculated between the performance scores of the Ss wthin each of the groups and are presented in Table 4. None of these correlations are significant (13). Thus the lack of correlation in performance could not be attributed to differences in involvement or in consistency of involvement, as measured by the eosinophil responsivity score.

Secondly, the obtained lack of correlation might be due to interaction between techniques and counterbalanced order. For example, the relationship between stress performance scores in the group which experienced failure first might be opposite in direction to that of the group which experienced distraction stress first—possibly a function of attitudes developed in the initial session. The correlation for the whole group under such conditions would

TABLE 4

RHO CORRELATIONS* BETWEEN STRESS PERFORMANCES WITHIN THREE GROUPS OF *S*s AS DEFINED BY EOSINOPHIL RESPONSIVITY

| Group | *N* | Failure stress scores vs.<br>Distraction overall stress scores | Distraction stress scores | Distraction free scores |
|---|---|---|---|---|
| Consistent Responders | 14 | —.30 | —.28 | —.19 |
| Consistent Non-Responders | 15 | +.12 | +.12 | —.26 |
| Inconsistent Responders | 20 | +.13 | +.05 | +.12 |

*For *N*s of 14, 15, and 20, rhos of .54, .52, and .45, respectively, are needed to attain the .05 level of significance (13).

probably be near zero. To examine this question the relationships between stress scores were tested separately in each group. The results appear in Table 5. None of the correlations are significantly different from zero. However, it should be noted that the correlations in the group which had failure stress first are positive, while correlations in the group which had distraction stress first are negative. A test was made for differences between the correlations for the two groups. No significant differences are found although the *p* values are between the .20 and the .10 levels of significance.

TABLE 5

PRODUCT-MOMENT CORRELATIONS* BETWEEN STRESS PERFORMANCES WITHIN THE FAIL-DISTRACT ORDER GROUP AND DISTRACT-FAIL ORDER GROUP

| Order | *N* | Failure stress scores vs.<br>Distraction overall stress scores | Distraction stress scores | Distraction free stress scores |
|---|---|---|---|---|
| Fail-Distract order ($N = 30$) | 30 | +.20 | +.05 | +.19 |
| Distract-Fail order ($N = 27$) | 27 | —.26 | —.34 | —.20 |
| Significance level of difference between two** correlations | | $p = .10$ | $p = .15$ | $p = .16$ |

*For *N*s of 27 and 30, correlations of .38 and .36 respectively, are necessary to attain the .05 level of confidence (5).

**Evaluated by Fisher's *r* to *z* transformation (5).

## D. DISCUSSION

No significant relationships between the stress performance scores derived from the two different experimental conditions are found in this experiment. These results are consistent with those reported by Stopol (17), and are not attributable to differences in involvement, as measured by the eosinophil responsivity score. Tests for significance between the oppositely signed correlations of the two order of experimentation groups, while they do not reach an acceptable level of confidence ($.20 < p > .10$) suggest that consideration might be given to the effect of different sequences of experimentation.

The results do not appear to substantiate the concept of a general factor of resistance to stressful conditions. Nevertheless, it could be justifiably argued that an estimate of a general factor, on the basis of behavior in one or two situations, might be attenuated by the presence of variables unique to these situations. If, however, behavior in many stressful situations were measured, a better estimate of any general factor might be obtained. With such an estimate prediction of specific behaviors might be possible.

Support for this line of reasoning is found in an experiment by Feffer and Phillips (7), who used the same group of *S*s and the same performance scores reported in the present experiment. Assuming that the tasks and problems encountered in everyday living could be classified as "stresses," these investigators rated each *S*'s adequacy in various areas, e.g., occupation-education, heterosexual relationships, and social relationships. *S*'s total scores were then related to performance under distraction stress and failure stress conditions. Despite the fact that these specific behaviors, i.e., performances under the two stress conditions, are unrelated, both were related to what could be considered a more general measure of resistance to stressful conditions, i.e., social adequacy.

While these relationships were significant, they did not attain a high level of confidence. To achieve substantial predictability in any one stress situation, it would appear necessary to also take into consideration those psychological variables which may be unique to the particular experimental conditions employed.

## E. SUMMARY

From the assumption of a general factor of resistance to stressful conditions, it can be inferred that performances on the same psychomotor task under two different stressful conditions should be related. To investigate this hypothesis, 57 normal, adult, male industrial workers were given, in counterbalanced order, a hand-steadiness test under distraction stress and failure stress conditions. An attempt was made to control for involvement in the two situations by measurement of the circulating eosinophils in the blood stream.

*S*s showed a significant and approximately equal responsivity in terms of eosinophil levels in both situations. Prestress levels of performance were also significantly related. However, no significant correlations were found between

the stress performance scores (corrected for prestress levels). Analysis of relationships within sub-groups defined by (*a*) order of experimentation, and (*b*) eosinophil responsivity, also failed to attain significance.

It was suggested that different psychological variables are pertinent to the two stress situations and that these variables should be taken into consideration if high predictability is to be attained.

## REFERENCES

1. CARLSON, V. R., & LAZARUS, R. S. A repetition of Meyer Williams' study of intellectual control under stress and associated Rorschach factors. *J. Consult. Psychol.,* 1953, **17**, 247-253.
2. DEESE, J., & LAZARUS, R. S. The effects of psychological stress upon perceptual motor performance. *Human Resources Research Center Bull.,* 1952, **52**, 52-10, 1-15.
3. DISCOMBE, G. Criteria of eosinophils. *Lancet,* 1946, **250**, 195-196.
4. DUNLAP, K. Improved forms of steadiness tester and tapping plate. *J. Exp. Psychol.,* 1921, **4**, 430-433.
5. EDWARDS, A. L. Experimental Design in Psychological Research. New York: Rinehart, 1950.
6. ERIKSEN, C. W., LAZARUS, R. S., & STRANGE, J. R. Stress and its personality correlates. *J. Personal.,* 1952, **20**, 277-286.
7. FEFFER, M., & PHILLIPS, L. Social attainment and performance under stress. *J. Personal.,* 1953, **22**, 284-297.
8. FORSHAM, P. H., THORN, G. W., PRUNTY, F. J., & HILLS, A. G. Clinical studies with pituitary adrenocorticotropin. *J. Clin. Endocrinol.,* 1948, **8**, 15-66.
9. HUNT, J. McV., & COFER, C. N. Psychological deficit. *In* J. McV. Hunt (*Ed.*), *Personality and the Behavior Disorders,* Vol. II. New York: Ronald Press, 1944.
10. LAZARUS, R. S., DEESE, J., & OSLER, S. J. The effects of psychological stress upon performance. *Psychol. Bull.,* 1952, **49**, 293-317.
11. MALMO, R. B., SHAGASS, C., BELANGER, D. J., & SMITH, A. A. Motor control in psychiatric patients under experimental stress. *J. Abn. & Soc. Psychol.,* 1951, **46**, 539-547.
12. MALMO, R. B., SHAGASS, C., & DAVIS, J. F. Electromyographic studies of muscular tension in psychiatric patients under stress. *J. Clin. & Exper. Psychopath.,* 1951, **12**, 45-66.
13. OLDS, E. G. The 5% significance levels for sums of squares of rank differences and a correction. *Ann. Math. Statis.,* 1949, **20**, 117-118.
14. PHILLIPS, L., & COWITZ, B. Social attainment and reactions to stress. *J. Personal.,* 1953, **22**, 270-283.
15. ROSENZWEIG, S. An outline of frustration theory. *In* J. McV. Hunt (*Ed.*), *Personality and the Behavior Disorders.* New York: Ronald Press, 1944.
16. SELYE, H. Annual Report on Stress. Montreal: Rinehart, 1950.
17. STOPOL, M. S. An experimental investigation of the consistency of stress tolerance and related Rorschach factors. Unpublished dissertation on file at the Columbia University Library, 1951.
18. WILLIAMS, M. An experimental study of intellectual control under stress and associated Rorschach factors. *J. Consult. Psychol.,* 1947, **11**, 21-29.

*Worcester State Hospital*
*Worcester, Massachusetts*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 467-474.

# DIMENSIONS OF COLOR VISION*

*Department of Psychology, University of Stockholm*

---

GOSTA EKMAN[1]

---

## A. PROBLEM

There have been many theories concerned with color vision, most of them falling into two or three main categories. No one theory seems to account for all the facts of the field—the laws of color mixing, the types of color blindness, etc.

In recent years physiologists have obtained very promising results with methods of nerve fibre analysis. The work of Granit on animals is well known also to psychologists. It clearly demonstrates the existence (in some species) of at least four types (or groups) of receptors, approximately corresponding to the four "unique" colors. These physiological findings will be considered in a subsequent section.

This paper describes a new psychological approach to the problem of primary dimensions of color vision. It is psychological with regard to both problem and method, but an effort will be made to link the results up with the physiological findings.

## B. METHOD

The method of *similarity analysis* was developed by the present writer for studies of the dimensionality of experience. Very briefly, its main features may be outlined as follows: Stimuli are presented two at a time. The subject is instructed to rate, on a suitable scale, the degree of subjective similarity between the stimuli. With $n$ stimuli, a similarity matrix of the order $n$ x $n$ will thus be obtained. The entries of this matrix may be individual data or group data.

The theory of the method is based upon the reasonable assumption that the degree of perceived similarity is a function of the degree of overlap between those primary experiences (sensations, emotions) which are evoked by the stimuli. Under certain conditions the methods of factor analysis may be directly applied to the similarity matrix. The entries of the resulting factor

---

*Received in the Editorial Office on July 5, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was made possible by a research grant from Magnus Bergvalls Stiftelse. The writer is indebted to Mr. M. Björkman, Mr. S. Borg, and Mr. T. Künnapas for valuable assistance in laboratory and computational work.

matrix indicate the relative contribution of any primary experience to the complex experience under consideration.

## C. THE EXPERIMENT

The subjects were looking at a screen from a distance of about 250 cm. There were two circular windows in the screen, 15 mm in diameter and 10 mm apart. They were covered with opaque glass and lighted from two projectors behind the screen. Different color filters could be inserted in the projectors. The experiments were conducted in a faintly lighted room.

Fourteen color filters were used, transmitting light of wave lengths 434 m$\mu$ to 674 m$\mu$ (see Table 1). The half-band width of these filters is about 12 m$\mu$. Every stimulus was combined with every other stimulus in a random order. The sequence of stimulus combinations was rotated among subjects. Each combination was presented for about 20 seconds.

TABLE 1
SIMILARITY MATRIX

| Wave length | 434 | 445 | 465 | 472 | 490 | 504 | 537 | 555 | 584 | 600 | 610 | 628 | 651 | 674 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 434 |  | .86 | .42 | .42 | .18 | .06 | .07 | .04 | .02 | .07 | .09 | .12 | .13 | .16 |
| 445 | .86 |  | .50 | .44 | .22 | .09 | .07 | .07 | .02 | .04 | .07 | .11 | .13 | .14 |
| 465 | .42 | .50 |  | .81 | .47 | .17 | .10 | .08 | .02 | .01 | .02 | .01 | .05 | .03 |
| 472 | .42 | .44 | .81 |  | .54 | .25 | .10 | .09 | .02 | .01 | .00 | .01 | .02 | .04 |
| 490 | .18 | .22 | .47 | .54 |  | .61 | .31 | .26 | .07 | .02 | .02 | .01 | .02 | .00 |
| 504 | .06 | .09 | .17 | .25 | .61 |  | .62 | .45 | .14 | .08 | .02 | .02 | .02 | .01 |
| 537 | .07 | .07 | .10 | .10 | .31 | .62 |  | .73 | .22 | .14 | .05 | .02 | .02 | .00 |
| 555 | .04 | .07 | .08 | .09 | .26 | .45 | .73 |  | .33 | .19 | .04 | .03 | .02 | .02 |
| 584 | .02 | .02 | .02 | .02 | .07 | .14 | .22 | .33 |  | .58 | .37 | .27 | .20 | .23 |
| 600 | .07 | .04 | .01 | .01 | .02 | .08 | .14 | .19 | .58 |  | .74 | .50 | .41 | .28 |
| 610 | .09 | .07 | .02 | .00 | .02 | .02 | .05 | .04 | .37 | .74 |  | .76 | .62 | .55 |
| 628 | .12 | .11 | .01 | .01 | .01 | .02 | .02 | .03 | .27 | .50 | .76 |  | .85 | .68 |
| 651 | .13 | .13 | .05 | .02 | .02 | .02 | .02 | .02 | .20 | .41 | .62 | .85 |  | .76 |
| 674 | .16 | .14 | .03 | .04 | .00 | .01 | .00 | .02 | .23 | .28 | .55 | .68 | .76 |  |

A number of preliminary trials were given to make the subjects acquainted with the situation. Forms had been prepared for the 91 paired comparisons. The subjects were instructed to rate the degree of "qualitative similarity" on a scale with five steps, ranging from 0 ("no similarity at all") to 4 ("identity").

The subjects were 31 students with normal color vision. All of them had some previous training in psychological laboratory work.

## D. Analysis of Data

The method of similarity analysis is applicable to individual data. This, in general, would require rather intensive experimentation with the single subject. In this case the group data were analyzed.

The similarity scores were averaged and transformed to a scale ranging from 0 to 1. These data are entered in Table 1.

The similarity matrix was factored by the group centroid method described by Thurstone (2). The orthogonal factor matrix is given in Table 2. The question of how far to continue the analysis was solved in the following way: The sample of 31 subjects was divided in two random halves of 15 and 16 individuals respectively. Similarity matrices were computed for both of these subgroups, and the difference matrix was obtained. The distribution of these differences is shown in Table 3. Factorization was continued until the distri-

TABLE 2

Orthogonal Factor Matrix

| Wave length | I | II | III | IV | V | $h^2$ |
|---|---|---|---|---|---|---|
| 434 | .131 | .676 | —.457 | .320 | .002 | .859 |
| 445 | .114 | .720 | —.426 | .384 | —.015 | .861 |
| 465 | .031 | .780 | —.081 | —.284 | .095 | .706 |
| 472 | .022 | .803 | —.019 | —.334 | .079 | .764 |
| 490 | .031 | .644 | .370 | —.227 | —.033 | .605 |
| 504 | .065 | .437 | .617 | .042 | —.130 | .595 |
| 537 | .100 | .303 | .672 | .378 | —.131 | .713 |
| 555 | .140 | .230 | .647 | .422 | —.066 | .674 |
| 584 | .497 | .023 | .288 | .185 | .353 | .489 |
| 600 | .724 | —.014 | .135 | .100 | .491 | .794 |
| 610 | .847 | —.028 | —.044 | —.064 | .151 | .747 |
| 628 | .871 | —.015 | —.111 | —.082 | —.213 | .823 |
| 651 | .822 | .011 | —.134 | —.100 | —.399 | .863 |
| 674 | .726 | .023 | —.136 | —.041 | —.385 | .696 |

TABLE 3

Distribution of Sample Differences and of Fifth Factor Residuals

| Interval | Frequency of Differences | Residuals |
|---|---|---|
| .000 — .029 | 45 | 63 |
| .030 — .059 | 24 | 18 |
| .060 — .089 | 11 | 8 |
| .090 — .119 | 4 | 2 |
| .120 — .149 | 2 | |
| .150 — .179 | 3 | |
| .180 — .209 | 2 | |

bution of residuals was well within the distribution of differences, as may also be seen in Table 3. Five factors were extracted.

It will be noted that in similarity analysis a sample of 31 individuals is comparable, from the point of view of statistical reliability, to a very much larger sample in correlation analysis with factorial methods. Sampling fluctuations, of course, are very much smaller in these arithmetical means than in corresponding coefficients of correlation. This is partly a consequence of the fact that the different observers agree rather well on their ratings.

### E. ALTERNATIVE SOLUTIONS

The factor matrix was rotated to simple structure (see below). The transformation matrix is given in Table 4 and the rotated orthogonal factor matrix in Table 5.

TABLE 4
TRANSFORMATION MATRIX

| | V | B | G | Y | R |
|---|---|---|---|---|---|
| I | —.044 | .010 | .115 | .619 | .776 |
| II | .652 | .721 | .225 | —.055 | —.025 |
| III | —.487 | .178 | .839 | .063 | —.146 |
| IV | .581 | —.662 | .444 | .072 | —.158 |
| V | —.006 | .102 | —.186 | .778 | —.591 |

TABLE 5
ROTATED FACTOR MATRIX

| Wave length | V | B | G | Y | R |
|---|---|---|---|---|---|
| 434 | .913 | .128 | —.030 | .046 | .084 |
| 445 | .904 | .188 | —.009 | .020 | .080 |
| 465 | .385 | .746 | —.030 | .024 | .006 |
| 472 | .341 | .804 | .004 | .001 | .006 |
| 490 | .109 | .678 | .364 | —.035 | .009 |
| 504 | .012 | .381 | .666 | —.043 | .019 |
| 537 | .095 | .074 | .836 | .013 | —.010 |
| 555 | .086 | —.006 | .810 | .093 | —.019 |
| 584 | .003 | —.020 | .320 | .612 | .104 |
| 600 | .013 | —.004 | .146 | .846 | .235 |
| 610 | .004 | .028 | —.002 | .636 | .585 |
| 628 | .038 | .000 | .069 | .360 | .833 |
| 651 | .054 | .008 | .014 | .182 | .909 |
| 674 | .093 | —.021 | .028 | .137 | .818 |

With five factors there are 10 plots. Out of these, eight plots yield a very satisfying simple structure with orthogonal axes. Two of the diagrams in Figure 1 are representative of these eight plots. There is practically no margin for alternative solutions, if negative loadings are to be avoided, which seems to be a quite reasonable criterion.

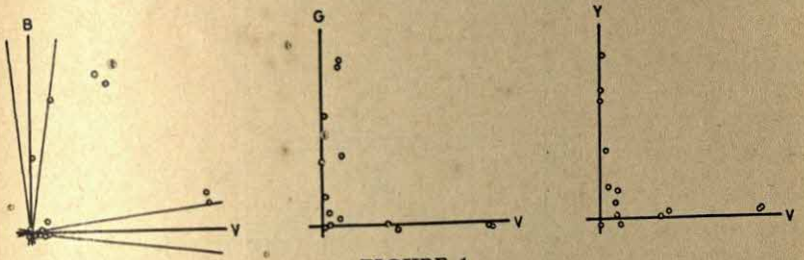


FIGURE 1

All the rotations were performed blindly up to the choice in two of the plots, which will now be described. One of them (B/V) is shown in the left part of Figure 1.

This diagram clearly indicates the possibility of obtaining simple structure by locating the axes at acute angles with each other. Obliqueness, of course, would have to be interpreted in terms of similarity between the primary sensations. At the present stage of similarity analysis, however, there is hardly any good reason for preferring the oblique solution to an orthogonal solution with some more small positive loadings.

The diagram indicates both of the extreme solutions for the factors under consideration. The compromise solution indicated by the bisecting vectors *V* and *B* is the one chosen for Tables 4 and 5. The margin of uncertainty is much the same in the plot of Y and *R*, where the compromise solution was chosen in the same way.

## F. Primary Color Factors

It will be remembered that the factor loadings found in our analysis indicate the relative contributions of the primary sensation factors to the complex sensation evoked by a certain physical stimulation.

This is one of the very few investigations where factor loadings can be plotted against a physical continuum. We shall take advantage of this fact and present the factors as curves over the base-line of wave length.

We start from the lower end of the spectrum. Our first factor, then, is represented in Figure 2 by two curves, indicating the two extreme solutions of the rotational problem. Whichever of the solutions we might choose between the two extremes, it is quite clear that this factor is in the violet end of the spectrum. It has been denoted by *V*.

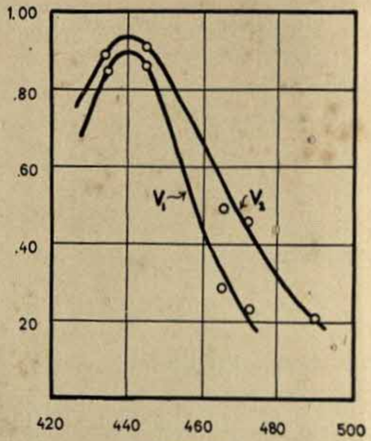


FIGURE 2

The area between the two curves in Figure 2 indicates the margin of uncertainty in the location of this factor. It is of the same order of magnitude in the location of the factors *B*, *Y*, and *R* (see below), i.e., it does not affect the interpretation of the factors. Subsequently, we present only the compromise solutions mentioned in the preceding section. All the five factors are shown in Figure 3.

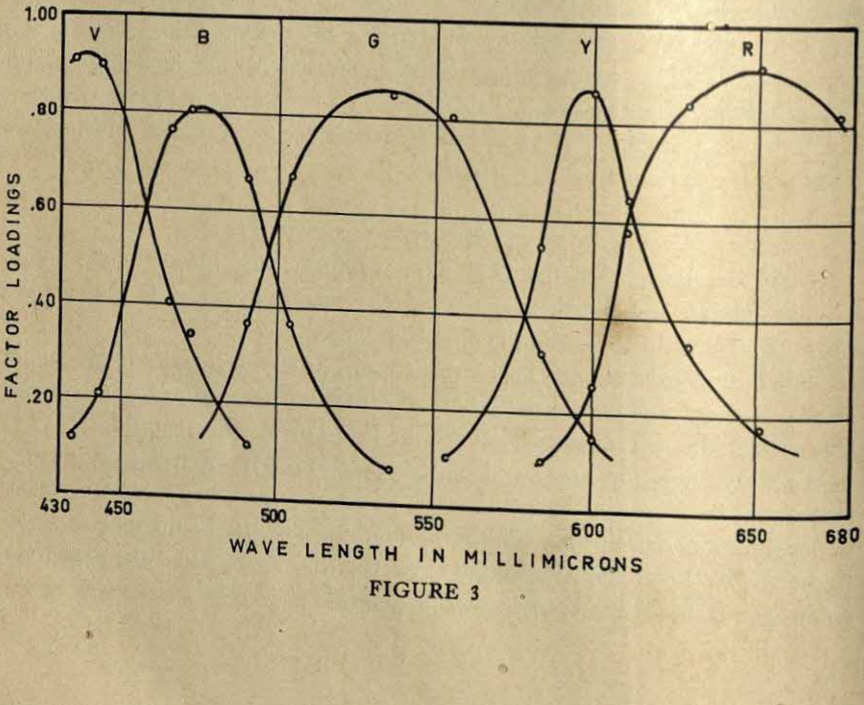


FIGURE 3

The second factor is located in the blue part of the spectrum. This factor is called *B*.

The third factor is located without any noticeable margin of uncertainty. It is in the broad green zone and will be called *G*.

The fourth factor is in the narrow yellow zone, and it is accordingly called *Y*.

The fifth and last factor is located in the red part of the spectrum, and the factor will be referred to as *R*.

These and no other factors were found in our analysis. It will be remembered that the subjects were instructed to rate "qualitative" similarity, which to most of them will have meant hue. This instruction was intentional. It was desired to rule out of the judgments, as far as possible, the attributes of luminosity and saturation. One reason for this condition was the difficulty, which might be caused by "quantitative" differences in the same dimension.

Accordingly, the factors have to be interpreted as factors in the perception of hue. Additional factors may account for other attributes of color, but they do not concern us now. Our analysis, then, would indicate that five primary sensations, in different combinations, are responsible for all the more or less complex sensations of hue which are evoked by light stimulation of different wave lengths.

## G. Discussion of Results

A comparison of our five factors with the physiological findings of Granit (1) yields the following results:

1. The blue receptor is located at about the same position as a composite curve that might be drawn from our *V* and *B* curves. This may or may not be due to incomplete differentiation of the physiological analysis.

2. The green curves agree very well.

3. The yellow curves agree fairly well, but our curve is located somewhat more to the right (cf. below).

4. Our *R* curve is located considerably more to the right than the red receptors. This is probably due to the difference in the meaning of the scales on the vertical axes. As we move to the right from about 600 m$\mu$, the "red" is certainly becoming rapidly darker, which may be taken as an indication of the quickly decreasing excitability of the receptor which may be at work. Our results, however, do not show this decrease, because the ratings were restricted to qualitative similarity, and the *hue* remains much the same. Furthermore, the factor loadings indicate the *relative* contribution of the red component of sensation, and this, of course, increases with diminishing contributions of the

yellow component. So far there would be no difficulty in connecting the $R$ factor with the red receptor.

5. The general shape of the curves agrees remarkably well. On the whole they come pretty close to the normal curve.

In conclusion, it may be stated that there would seem to be some good reasons for connnecting the factors of our analysis with Granit's receptor groups. This, of course, would not mean anything like identification. One of the series is psychological, sensational, and the other is physiological; and it is certainly far beyond the scope of this paper to bridge this philosophical gap. The connection would be entirely in terms of the common element: the common physical continuum.

The possibility of a connection of this type might be quite important. It would mean that receptor analysis could serve as a guide for psychological dimensional research with this or any comparable method, and that the opposite would be equally true.

It would signify for instance, that the dimensionality of olfaction could be approached along either or both of these ways. It is the latter course that ought to be adopted.

## REFERENCES

1. GRANIT, R. Sensory Mechanisms of the Retina. Oxford: Univ. Press, 1947.
2. THURSTONE, L. L. Multiple-Factor Analysis. Chicago: Univ. Press, 1947.

*Department of Psychology*
*University of Stockholm*
*Teknologgatan 8*
*Stockholm Va, Sweden*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 475-481.

# RESPONSE BIAS IN AN UNSTRUCTURED QUESTIONNAIRE*

*Department of Psychology, Northwestern University*

IRWIN A. BERG AND GERALD M. RAPAPORT

## A. THE PROBLEM

In stimulus situations which permit a variety of reactions subjects often exhibit a preference for certain responses as opposed to others. As a result, the distribution of human responses is typically quite different from what would be expected on the basis of statistical probability. Manifestation of this bias in test situations has been labelled "response set" by Cronbach (5). He describes a number of sets such as *acquiescence,* or the tendency to prefer options as "true" or "yes," and *evasiveness,* or the tendency to select such options as "indifferent," "uncertain," etc. It has been observed that these biased responses are most likely to appear in ambiguous or unstructured test situations (5) or when the test items provide no direct or indirect clues which would permit the subject to select a given option (10). The extent to which such biases appear when structure is virtually absent is not clear, although isolated studies have demonstrated that such sets exist. Goodfellow (8), to give one example, found a tendency for a flipped coin to be called "heads" 80 per cent of the time on the first toss.

## B. PURPOSE

A number of writers (1, 2, 3, 5, 7) have suggested on the basis of some evidence and some speculation that these biased response patterns are probably reflections of personality characteristics. While this interpretation appears feasible, it is important to know to what extent biased responses will appear in a maximally unstructured and simple test situation. If such response sets do not appear significantly under conditions most favorable for eliciting them, serious doubt would be cast upon their usefulness as possible measures of personality. Certainly their value for such purposes would be quite limited. With this question in mind, an imaginary questionnaire was prepared and administered to groups of college students. It was hypothesized that without item content to provide a basis for directing the subjects' responses, any tendencies toward response sets should appear in the groups.

---

*Received in the Editorial Office on July 6, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

## C. SUBJECTS AND PROCEDURE

An imaginary questionnaire was administered to 374 students (188 males, 186 females) in undergraduate college classes. The questionnaire supplied to each subject was a mimeographed sheet with instructions requesting the subject to choose one option from a group of options offered for nine imaginary questions. Three of the items were two-option choices, as *heads* or *tails;* three were three-option choices, as *yes, uncertain, no;* three were four-option choices, as *1, 2, 3, 4.* See Table 1 for the order and types of options used. For 171 subjects (77 males and 94 females) the questionnaire was administered in regular order, i.e., *heads, tails; true, false,* etc. To provide a check on position preferences, this order was reversed for 203 subjects (111 males and 92 females), i.e., *tails, heads; false, true,* etc.

When each subject had been supplied with a questionnaire form, the subjects were asked to read the following instructions at the top of the sheet while the examiner read them aloud:

> This is an experiment to determine how well you can guess the answers to questions when you do not know what the question is. The experimenter will call off the Roman numeral for each question, and he will imagine the correct answer. You are merely to circle one of the alternatives in each case. Just select whatever answer enters your mind without trying to think heavily about it. *Please do not begin until the signal is given. Remember,* circle one alternative only.

The examiner would then call out, "Roman numeral I, circle either *heads* or *tails*" and so on, repeating the Roman numerals and the options in each case. Male and female graduate students, who knew that the study involved "set" in an unstructured situation but who did not know anything of the hypothesized relationships, were used as examiners. They did not, of course, attempt to "imagine" any correct answer. When the data were gathered and tabulated, chi square values were computed for the distribution of responses to each option.

## D. RESULTS AND DISCUSSION

An examination of Tables 1 and 2 will offer abundant evidence that in an ambiguous or unstructured situation the distribution of human responses does not follow the laws of statistical probability. For some items the response bias is so marked that chi square values of greater than 80 are obtained, attesting to an exceedingly high level of statistical significance. These findings are not peculiar to the present situation, for similar bias may be

TABLE 1

DISTRIBUTION OF RESPONSES TO THE IMAGINARY QUESTIONNAIRE: REGULAR FORM

| | | Male | | Female | | Total *N* | |
|---|---|---|---|---|---|---|---|
| Item | Option | *N* | $X^2$ | *N* | $X^2$ | *N* | $X^2$ |
| I | Heads | 39 | | 72 | | 111 | |
| | Tails | 38 | 0 | 22 | 25.5** | 60 | 14.6** |
| II | 1 | 7 | | 7 | | 14 | |
| | 2 | 21 | | 20 | | 41 | |
| | 3 | 40 | | 52 | | 92 | |
| | 4 | 9 | 35.** | 15 | 49.** | 24 | 84.0** |
| III | Yes | 35 | | 47 | | 82 | |
| | Uncertain | 21 | | 22 | | 43 | |
| | No | 21 | 5.0 | 25 | 11.8** | 46 | 16.5** |
| IV | X | 46 | | 60 | | 106 | |
| | Y | 31 | 2.5 | 34 | 4.5* | 65 | 9.3** |
| V | Very Satisfied | 9 | | 13 | | 22 | |
| | Satisfied | 43 | | 50 | | 93 | |
| | Dissatisfied | 16 | | 24 | | 40 | |
| | Very Dissatisfied | 9 | 40.** | 7 | 46.** | 16 | 85.0** |
| VI | First | 31 | | 31 | | 62 | |
| | Second | 24 | | 35 | | 59 | |
| | Third | 22 | 1.7 | 28 | .7 | 50 | 1.3 |
| VII | True | 49 | | 69 | | 118 | |
| | False | 28 | 5.2* | 25 | 19.6** | 53 | 23.9** |
| VIII | A | 13 | | 18 | | 31 | |
| | B | 34 | | 30 | | 64 | |
| | C | 18 | | 29 | | 47 | |
| | D | 12 | 16.** | 17 | 6.2 | 29 | 18.0** |
| IX | Agree | 32 | | 40 | | 72 | |
| | Indifferent | 20 | | 32 | | 52 | |
| | Disagree | 25 | 2.8 | 22 | 5.1 | 47 | 6.1* |

*5% level of significance. **1% level of significance.

found under other free conditions. For example, in a large undergraduate course with which the senior author of the present study is associated, the students are required early in the course to prepare four multiple-choice items each week on the lecture and reading materials. The only instructions are that each item be placed on a standard 4 x 6 inch card, have four options, and the number of the correct option and its source be placed on the bottom of each card. The correct answer could, of course, be placed in any of the four positions since no directions concerning answer placement were given. Yet when 1000 consecutive cards obtained from the first assignments given to 400 students were tabulated for location of the correct answers, the distribution was far from the theoretical expectation of 250 correct answers assigned to each of the four options. For the 1000 items (an average of

TABLE 2

DISTRIBUTION OF RESPONSES TO THE IMAGINARY QUESTIONNAIRE: INVERTED FORM

| Item | Option | Male $N$ | Male $X^2$ | Female $N$ | Female $X^2$ | Total $N$ $N$ | Total $N$ $X^2$ |
|---|---|---|---|---|---|---|---|
| I | Tails | 53 | | 39 | | 92 | |
| | Heads | 58 | .14 | 53 | 2.00 | 111 | 1.49 |
| II | 4 | 20 | | 11 | | 31 | |
| | 3 | 45 | | 44 | | 89 | |
| | 2 | 27 | | 33 | | 60 | |
| | 1 | 19 | 15.6** | 4 | 30.5** | 23 | 53.1** |
| III | No | 25 | | 24 | | 49 | |
| | Uncertain | 25 | | 20 | | 45 | |
| | Yes | 61 | 23.3** | 48 | 14.8** | 109 | 37.7** |
| IV | Y | 49 | | 35 | | 84 | |
| | X | 61 | 1.1 | 57 | 4.8* | 118 | 5.4* |
| V | Very Dissatisfied | 13 | | 12 | | 25 | |
| | Dissatisfied | 32 | | 23 | | 55 | |
| | Satisfied | 54 | | 44 | | 98 | |
| | Very Satisfied | 12 | 42.1** | 13 | 28.8** | 25 | 70.2** |
| VI | Third | 34 | | 14 | | 48 | |
| | Second | 33 | | 34 | | 67 | |
| | First | 44 | 2.0 | 44 | 15.1** | 88 | 11.7** |
| VII | False | 47 | | 31 | | 78 | |
| | True | 64 | 2.3 | 61 | 9.1** | 125 | 10.4** |
| VIII | D | 16 | | 11 | | 27 | |
| | C | 27 | | 28 | | 55 | |
| | B | 42 | | 36 | | 78 | |
| | A | 26 | 12.4** | 17 | 16.2** | 43 | 27.1** |
| IX | Disagree | 30 | | 20 | | 50 | |
| | Indifferent | 31 | | 31 | | 62 | |
| | Agree | 50 | 6.8* | 41 | 7.1* | 91 | 14.5** |

*5% level of significance. **1% level of significance.

2.5 items from each of 400 students) the correct answers were located as follows: option 1, 54; option 2, 237; option 3, 571; option 4, 138. For significance at the one per cent level a chi square value of only 11.3 would be required for these data. However, the actual chi square value for this distribution of answer position reaches the fantastic level of 616.7, a convincing demonstration of bias in a free situation.

Thus in a situation which is about as unstructured as it is feasible to arrange, response bias undeniably occurs. The question of whether this bias is due primarily to the order or position of the options or, perhaps, to some predilection for option content, minimal as the latter may be, can be assessed by scrutinizing Tables 1 and 2. The four-choice options in Table 1, Roman Numerals II, V, VIII may be taken as a case in point. In

Item II, the bias occurs at No. 3, the third option, while for V and VIII the bias occurs at the second option, "satisfied" and "B" respectively. Hence the bias shifts from the third to the second position with changes in item content. If this bias is chiefly a function of option content, we should expect that reversing the order of the options would then reveal a bias favoring a different position but for the same content. The results of the inverted form of the imaginary questionnaire given in Table 2 show that such was the case. Although Item II has the No. 3 located in second position, the bias continues to be in favor of the No. 3, not the third position. For Items V and VIII in which the second position had originally been preferred, the bias is now found at the third position, the new location of the same option content.

Accordingly, it seems reasonable to assert that bias in the present situation is chiefly, though not exclusively, a function of option content, minimal though it may be. Position and other preferences very likely exist in the present situation, but not as the primary cause of bias. It should be noted in this connection that the chi square values, though remaining significant in most instances, do tend to decrease when the regular and inverted forms are compared (see Tables 1 and 2). If option content alone were responsible for the obtained bias, we should expect the level of bias to be fairly constant in all instances when the regular and inverted forms of the imaginary questionnaire were compared. Since such was not the case, we may infer that other factors exerted some influence. Indeed, there is some evidence which indicates that position preference may be a factor in bias. For example, in an unpublished study by Rubin which Cronbach (6) quotes, a preference was found for the first option given in a pitch discrimination test and similar situations. But in the present situation, position preference is not the major factor in biased responses.

Though also secondary to content, other influences beside position sets probably operate in determining bias. In Tables 1 and 2, for example, the bias is occasionally confined to the males or the females alone, as in Items I and VIII of Table 1. It is believed that these sex differences in bias were probably the result of the sex and rôles of the examiners. No control over the examiner as a variable was instituted, unfortunately, because the possible influence of young, attractive examiners upon the subjects was not anticipated. Further studies are being planned with controls over this variable incorporated in the design.

The biased responses which favor option content in the present study are probably the product of cultural factors. In our culture, for example, it is characteristic, as a popular song puts it, to "accentuate the positive" and that we respond pleasantly in social situations akin to the "hello-goodbye" pattern described by Hathaway (9). The "yes" man traditionally moves in our culture with less friction than the "no" man, since we value such conformity positively. It is thus not surprising that we find significant preferences for such option content as "yes," "satisfied," "true," and "agree" in the present data. Such responses would be in the direction of established custom. Other cultures or sub-cultures, of course, may value the neutral or negative response, and in such instances we should expect biased responses in the direction of neutrality or negativism.

While the present discussion has been confined to sets to respond in a given direction, it should be noted that the deviate responses, i.e., those responses which are different from the conventional ones, are believed to be most significant for future research. Earlier in the present study, it was noted that response sets have been considered by various writers to be reflections of personality characteristics. While no data may be appropriately presented at this time, some findings of current investigations may be mentioned briefly in this connection. The important factor in biased responses appears to be centered, not so much in the direction of bias itself, but rather in responses which consistently depart from the bias. That is, subjects who often choose "no" or "dissatisfied" when the popular response is "yes" or "satisfied" have been found to be deviate in personality as well as in their responses to an unstructured test. Using a relatively unstructured nonverbal test (4), for example, Barnes (1) was able to construct useful clinical scales from deviate responses which significantly differentiated between a group of normal subjects and a group of subjects undergoing psychiatric treatment. Other unpublished evidence indicates similarly that preferences which rather consistently go counter to the modal preference are associated with personality deviation. Perhaps a method such as this, so simple as not to prompt unjustified generalizations beyond the logical boundaries of the specific individual-group relation thus isolated, warrants serious consideration as a parsimonious way of identifying operationally and meaningfully that all-too elusive concept, the deviate personality.

## E. SUMMARY

Response sets have been asserted to appear more frequently in unstructured or ambiguous test situations, and it is believed that these response tendencies are related to personality characteristics. To be useful as possible measures of personality, it should be demonstrated that response biases

would appear significantly in situations specifically designed to elicit them. Accordingly, a questionnaire form which had no actual questions but which required the subjects to imagine the correct answer when various options such as *true, false, yes, uncertain, no* were presented to them. Regular and inverted orders of options were employed in the questionnaires which were administered to 374 college students (188 males, 186 females). It was found that response biases appeared at a high level of statistical significance, reaching chi square values above 80 in some cases. Position preferences were not found to be the major factor in the observed biases. Rather the preferences were for certain culturally valued aspects of option content, "yes," "true," "agree," etc. Other cultures would, presumably, show different patterns of response bias.

## REFERENCES

1. BARNES, E. H. The relationship of biased test responses to psychopathology. Unpublished Ph.D. dissertation, Northwestern University, 1954.
2. BERG, I. A. The reliability of extreme position response sets in two tests. *J. of Psychol.*, 1953, **36**, 3-9.
3. BERG, I. A., & COLLIER, J. S. Personality and group differences in extreme response sets. *Educ. Psychol. Measm't*, 1953, **13**, 164-169.
4. BERG, I. A., & HUNT, W. A. The Perceptual Reaction Test. Evanston, Illinois, 1949.
5. CRONBACH, L. J. Response sets and test validity. *Educ. Psychol. Measm't*, 1946, **6**, 475-494.
6. ———. Further evidence on response sets and test designs. *Educ. Psychol. Measm't*, 1950, **10**, 3-31.
7. GAIER, E. L., LEE, MARILYN E., & MCQUITTY, L. L. Response patterns in a test of logical inference. *Educ. Psychol. Measm't*, 1953, **13**, 550-567.
8. GOODFELLOW, L. D. The human element in probability. *J. Gen. Psychol.*, 1940, **33**, 201-205.
9. HATHAWAY, S. R. Some considerations relative to nondirective counseling as therapy. *J. Clin. Psychol.*, 1948, **4**, 226-231.
10. RAPAPORT, G. M., & BERG, I. A. Response sets in a multiple-choice test. (Unpublished.)

*Department of Psychology*
*Northwestern University*
*Evanston, Illinois*

Published as a separate and in *The Journal of Psychology*, 1954, **38**, 483-493.

# NATURAL PREFERENCES OF THE CHICK FOR OBJECTS OF DIFFERENT COLORS*[1]

*Department of Psychology, University of Chicago*

ECKHARD H. HESS AND WALTER C. GOGEL[2]

## A. PROBLEM

The purpose of this experiment is to investigate the occurrence of natural color preferences in the chick. The preferences are defined as the tendency of the chick to peck more at small objects of one color than of another color, when the two kinds of differently colored objects are presented simultaneously.

It is clear, as shown by the work of Lashley (3) and others, that the chicken has chromatic vision. Also, it seems that the color vision of the chicken is similar to that of humans except that the chicken has an increased sensitivity to the red portion of the spectrum and a decreased sensitivity to the blue portion of the spectrum (4).

## B. APPARATUS

The apparatus consisted of a display box (10 inches long) containing two compartments in which chick mash was placed. Each compartment was 2.5 inches wide and 5 inches long. The display box was covered by a thin sheet of transparent Plexiglas so that the chicks could see the mash under the Plexiglas and could peck on the Plexiglas but could not strike the mash. The box was illuminated from above by a 60 watt frosted bulb. The mash (a chick starter mash) was colored by means of lead paint. Series of mash stimuli were prepared. One series (called a saturation series) varied from white to saturated red. Another series (called a lightness series) varied from white to dark grey. Also, chromatic stimuli other than red were prepared. The color of each of the mash stimuli was matched by the experimenters with the dull side of a color chip in the Ostwald Color Harmony Manual (1). This color match was used to specify the approximate reflectance of the achromatic stimuli and the approximate reflectance and purity

---

*Received in the Editorial Office on July 15, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This work was supported under contract with the Department of the Army, Office of the Surgeon General, on Medical Research Contract DA-49-007-MD-36.

[2] Now at Army Medical Research Laboratory, Fort Knox, Kentucky.

(saturation) of the chromatic stimuli (see Table 7). The color name for each color notation used in the tables and figures of this study is also given in Table 7.

## C. Procedure

One of the compartments of the display box was filled with a mash of a certain color and the other compartment was filled with a mash of a different color. Several chicks (usually three) at a time were allowed to peck for a specified interval of time (generally two minutes) on the Plexiglas covering the two kinds of mash. More than one chick was used at a time since a chick in isolation does not peck well (2). The two or three chicks which were used at a time will be called a "pair of chicks." The word "group" will refer to chicks of the same breed and the same age. The same group number appearing in several of the tables means that the results were obtained from the same group of chicks.

The experimenter counted the number of pecks given to each compartment of the display box by each pair of chicks. After a number of pairs of chicks had been used, the total amount of pecks given to each compartment and thus to each color for that particular comparison of colors was computed. The ratio of the total number of pecks on the two different colors gave what is called the pecking ratio of one color over that of the other. If the same chick or chicks from the same group, were tested more than once (on either the same or different days) for a particular pair of colors, the results were added together and the pecking ratio was then computed. The pecking ratio is interpreted as indicating the color pecking preference of the chick. During a test for the preference of one color over another, the one color was systematically sometimes on the right and sometimes on the left of the other color in the display box.

To find the change in the pecking preferences of the chick along either the white to red or the white to dark grey dimension, one of the mash stimuli was presented in comparison with each of the other stimuli of that series of stimuli. For example, in the white to red series, a white mash was presented simultaneously with a light pink mash and the pecking ratio on these two colors was determined. Similarly, the pecking ratio of the white mash in comparison with each of the other two stimuli in the series was found. The mash which was simultaneously presented with each of the other colors was sometimes a color other than white. When the pecking ratio resulting from the comparison of this constant stimulus with each of two or more stimuli (the variable stimulus) in the series is plotted against the change in the variable stimulus, the resulting curve is interpreted as showing a change in preference as a function of the change in the variable stimulus. The amount of pecking on the constant stimulus was always used in the denominator in computing the pecking ratio. The constant stimulus in comparison with itself was assigned a pecking ratio of "one."

## D. RESULTS AND DISCUSSION

Table 1 shows the summated results from using white, pink, and red stimuli. The chicks were from four days to approximately a month old at the time of the tests and had been fed only on tan chick feed. The notation $R_1$, $R_2$, and $R_4$ represents three red mash stimuli with the saturation increasing as the subscript increases. $R_1$ was a very light pink, $R_2$ was a more saturated pink, and $R_4$ was a saturated red. $W_1$ and $W_2$ stand for two white mash stimuli. $W_1$ had a lower reflectance than $W_2$. In Table 1

TABLE 1

PREFERENCE BEHAVIOR OF CHICKS WITH MASH WHICH DIFFERED IN LIGHTNESS AND SATURATION

These Chicks Had a Diet of Tan Mash Prior to the Tests. The number of pecks received by each color from *N* pairs of chicks is shown directly below the notation for each color.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | Group 1 | | | | |
| $W_1$ | $R_1$ | $W_1$ | $R_2$ | $W_1$ | $R_4$ | | |
| 146 | 121 | 281 | 72 | 157 | 20 | | |
| $N=20$ | | $N=32$ | | $N=25$ | | | |
| | | | Group 2 | | | | |
| $W_1$ | $R_1$ | $W_1$ | $R_2$ | $W_1$ | $R_4$ | $R_2$ | $R_4$ |
| 86 | 119 | 231 | 121 | 363 | 101 | 619 | 302 |
| $N=10$ | | $N=12$ | | $N=25$ | | $N=46$ | |
| | Group 3 | | | | | Group 5 | |
| $W_2$ | $R_2$ | $W_2$ | $R_4$ | $W_2$ | $R_4$ | $R_2$ | $R_4$ |
| 278 | 182 | 917 | 411 | 592 | 265 | 229 | 103 |
| $N=16$ | | $N=74$ | | $N=43$ | | $N=18$ | |
| | | | Group 4 | | | | |
| $W_2$ | $R_1$ | $W_2$ | $R_2$ | $W_2$ | $R_4$ | $R_2$ | $R_4$ |
| 151 | 331 | 347 | 802 | 275 | 248 | 1327 | 391 |
| $N=10$ | | $N=20$ | | $N=20$ | | $N=31$ | |
| | | | Group 6 | | | | |
| $W_2$ | $R_1$ | $R_1$ | $R_2$ | $R_1$ | $R_4$ | $R_2$ | $R_4$ |
| 54 | 153 | 549 | 246 | 581 | 130 | 141 | 17 |
| $N=10$ | | $N=20$ | | $N=19$ | | $N=10$ | |

and the following tables, the pairs of colors which were presented simultaneously to the chicks are shown at the top of each block in the table. The "N" in each block is the number of pairs of chicks which were used. The number of pecks received by each color from *N* pairs of chicks is shown directly below the notation for each color. For example, with Group 2, Table 1, 10 pairs of chicks were tested, one pair at a time on the simultaneously presented white ($W_1$) and pale pink mash ($R_1$). Adding the number of pecks given to each stimulus by each of the 10 pairs of chicks resulted in 86 pecks for the white and 119 pecks for the light pink mash.

Figure 1 shows the relation between the pecking ratios from Group 2

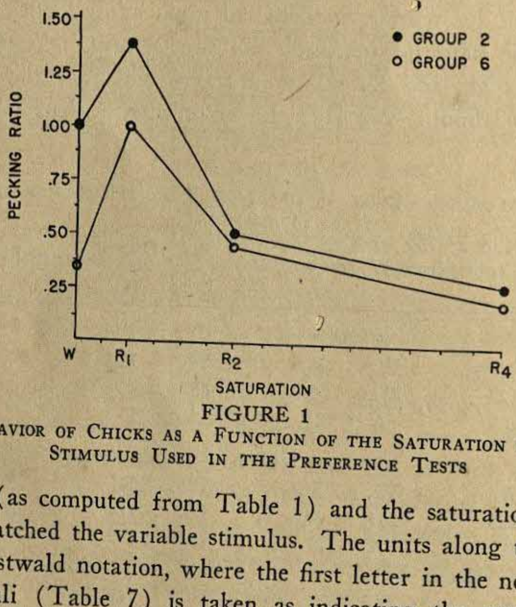


FIGURE 1
PREFERENCE BEHAVIOR OF CHICKS AS A FUNCTION OF THE SATURATION OF THE VARIABLE STIMULUS USED IN THE PREFERENCE TESTS

and Group 6 (as computed from Table 1) and the saturation of the color chips which matched the variable stimulus. The units along the abscissa are those of the Ostwald notation, where the first letter in the notation for this series of stimuli (Table 7) is taken as indicating the saturation of the variable stimulus. Thus there are four units of saturation between $R_1$ and $R_2$ and nine units of saturation between $R_2$ and $R_4$. For Group 2 the constant stimulus was $W_1$ and for Group 6 it was $R_1$. The greater the height along the ordinate, the larger the proportion of pecks which was received by the stimulus whose notation is on the abscissa. It is assumed that a stimulus color which has a higher ordinate value than another color would be preferred by the chicks over that color if the two were presented simultaneously. That this is correct is indicated, for example, by Group 4 in Table 1. The pecking ratio of $R_4$ over $W_2$ was less than that of $R_2$ over $W_2$, and $R_2$ was preferred over $R_4$.

Table 2 gives results from groups of chicks which were tested as to their preferences for several stimuli in the white to red series, where the tests were made before the several day-old chicks had any feeding experience. If the results in Tables 1 and 2 had been listed as pecking ratios they would indicate that the most preferred stimulus was in the desaturated portion of the saturation series. This is illustrated for part of Table 1 by Figure 1 in which the light pink stimulus was the most preferred color for Groups 2 and 6. The general similarity of Tables 1 and 2 in this respect suggests that the feeding experience on tan mash was not in general responsible for the preference behavior shown in Table 1.

TABLE 2

PREFERENCE BEHAVIOR OF CHICKS WITH MASH WHICH DIFFERED IN LIGHTNESS AND SATURATION

The Results Were Obtained Before the Chicks Were Put on Feed. The number of pecks received by each color from $N$ pairs of chicks is shown directly below the notation for each color.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Group 4 | | | | | |
| $W_2$ | $R_2$ | $W_2$ | $R_4$ | $R_2$ | $R_4$ | | |
| 158 | 732 | 240 | 176 | 420 | 162 | | |
| $N=14$ | | $N=19$ | | $N=16$ | | | |
| Group 3 | | Group 5 | | | Group 6 | | |
| $W_2$ | $R_4$ | $W_2$ | $R_2$ | $W_2$ | $R_2$ | $R_2$ | $R_4$ |
| 27 | 41 | 26 | 144 | 135 | 283 | 251 | 156 |
| $N=5$ | | $N=6$ | | $N=19$ | | $N=14$ | |
| | | Group 7 | | | | | |
| $W_1$ | $R_1$ | $W_2$ | $R_1$ | $R_1$ | $R_4$ | $R_2$ | $R_4$ |
| 93 | 202 | 87 | 288 | 352 | 253 | 269 | 55 |
| $N=18$ | | $N=10$ | | $N=20$ | | $N=10$ | |

The total amount of pecking of a pair of chicks in a test is a function of such factors as the activity level and motivation of the chicks. These factors are difficult to control and cannot be considered as constant throughout this study. The pecking ratio appears however to be a fairly reliable measure.

## 1. *The Effect of Lightness*

With the stimuli which varied from white to red, the lightness of the stimulus decreased as the saturation increased. It is therefore possible that the preference behavior of Table 1 and 2 was determined by lightness rather than by saturation. To consider this possibility, the pecking behavior of the chick along the white to dark grey dimension was investigated. The summarized results are given in Table 3. Also included are some comparisons between achromatic and chromatic stimuli. The results in the blocks which are marked with an asterisk were obtained before the chicks had feeding experience. The remaining data were obtained after the chicks had feeding experience on tan mash. In the case of Group 4, $R_4$ was compared with both chromatic and achromatic stimuli (Tables 2 and 3) before the chicks had eating experience, and $W_2$ was compared with both chromatic and achromatic stimuli (Tables 1 and 3) after the chicks had been fed on tan mash. Figure 2a shows the pecking ratios resulting from the comparison of each stimulus with $R_4$, plotted against the reflectance of the variable stimulus. Figure 2b shows the pecking ratios resulting from the comparison of each stimulus with $W_2$ plotted against the reflectance of the variable stimulus. The notation of the variable stimulus for each point is shown in the two figures. In both Figures 2a and 2b the solid line refers to the variable chromatic stimuli and the dashed line to the variable achromatic stimuli. The reflectance values of the chromatic stimuli must be regarded as only approximate since the illumination used in this experiment

TABLE 3

PREFERENCE BEHAVIOR OF CHICKS WITH MASH WHICH DIFFERED IN LIGHTNESS OR IN BOTH LIGHTNESS AND SATURATION

The results marked with an asterisk were obtained before the chicks were put on feed. The other results were obtained after the chicks had a diet of tan mash. The number of pecks received by each color from *N* pairs of chicks is shown directly below the notation for each color.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Group 3 | | | | Group 5 | | | |
| $W_2$ | MG | $R_2$ | MG | $W_2$ | DG | | |
| 338 | 252 | 363 | 824 | 271 | 142 | | |
| $N = 16$ | | $N = 34$ | | $N = 16$ | | | |
| | | Group 7 | | | | | |
| LG * | $W_2$ | LG * | $R_1$ | LG * | DG | | |
| 112 | 181 | 254 | 387 | 144 | 292 | | |
| N = 14 | | N = 23 | | N = 15 | | | |
| | | | Group 4 | | | | |
| $W_2$ | LG | $W_2$ | MG | $W_2$ | DG | $R_2$ | MG |
| 252 | 289 | 567 | 561 | 230 | 220 | 246 | 109 |
| $N = 14$ | | $N = 34$ | | $N = 10$ | | $N = 10$ | |
| $R_4$ * | LG | $R_4$ * | MG | $R_4$ * | DG | $R_4$ | MG |
| 82 | 95 | 146 | 329 | 98 | 67 | 525 | 843 |
| $N = 10$ | | $N = 20$ | | $N = 10$ | | $N = 20$ | |

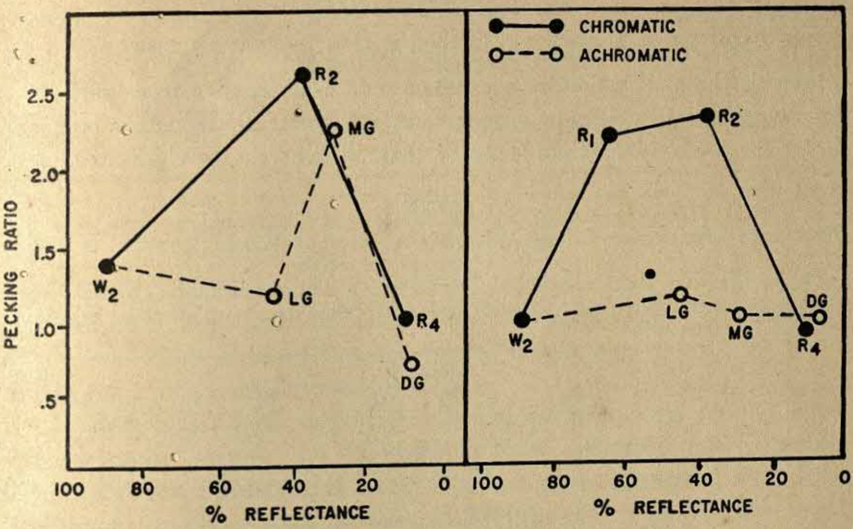


FIGURE 2

PREFERENCE BEHAVIOR OF CHICKS ALONG THE SATURATION AND LIGHTNESS DIMENSIONS WHEN THE VARIABLE STIMULUS IS EXPRESSED IN TERMS OF ITS REFLECTANCE

In Figure 2a the chicks were tested before they had eaten. In Figure 2b the chicks had a diet of tan mash prior to the tests.

was from a 60-watt frosted light bulb, whereas the reflectance values of the matching color chips had been determined under Illuminant C. In Figure 2a there are some differences between the chromatic and achromatic preference curves. There is a rise, however, in the achromatic preference curve at the medium grey. In Figure 2b this rise does not occur (using 34 pairs of chick in comparing $W_2$ with MG). The separation between the two curves of Figure 2b tends to be substantiated by the preference of Group 4 for $R_2$ over MG (Table 3). The differences between the achromatic and chromatic preference curves of Figure 2a and 2b suggest that saturation has an affect upon the chicks pecking preference in addition to any affect of lightness alone.

Not all of the groups were of the same breed. Groups 1, 2, and 3 were New Hampshires. Group 4 was Leghorns, Groups 5 and 6 were hybrids, and Group 7 was White Rocks. There seems to be some differences between the preferences of different groups of chicks. For example, with Group 3 (Table 3) the medium grey is preferred over $R_2$, while with Group 4 (Table 3) $R_2$ appears to be preferred over the medium grey. The reason for the differences between groups is not known.

### 2. *Different Hues*

The preferences of chicks which had been fed on tan mash were tested also with chromatic stimuli other than red. The results from such tests for two groups of chicks is given in Table 4 where *B* stands for blue and *G* for green. It seems from Table 4 that, in general, the chicks tended to prefer the lighter more desaturated color when presented with two colors which differed in lightness saturation and hue. This does not mean that the chicks have no hue preferences.

TABLE 4

PREFERENCE BEHAVIOR OF CHICKS WITH MASH, MOST OF WHICH DIFFERED IN LIGHTNESS, SATURATION AND HUE

The results were obtained after the chicks had a diet of tan mash. The number of pecks reecived by each color from *N* pairs of chicks is shown directly below the notation for each color.

| | | | | | |
|---|---|---|---|---|---|
| | | Group 3 | | | |
| $G_2$ | $B_3$ | $B_2$ | $W_2$ | $B_3$ | $R_4$ |
| 220 | 175 | 94 | 132 | 76 | 69 |
| $N = 23$ | | $N = 10$ | | $N = 10$ | |
| Group 6 | | | | Group 3 | |
| $R_1$ | $G_3$ | $R_4$ | $G_3$ | $G_4$ | $R_4$ |
| 892 | 73 | 515 | 168 | 41 | 96 |
| $N = 20$ | | $N = 30$ | | $N = 10$ | |

### 3. *The Naturalness of the Preferences*

The chicks in this experiment which were tested before they had eating experience were not without visual and pecking experience at the time of the tests. The question occurs as to the importance of this pre-feeding visual and pecking experience in determining the preference behavior. In a preliminary experiment, it had been found (Group *A,* Table 5) that Leghorn chicks preferred a tan over a blue-green mash when the chicks were kept in a tan environment and were tested before being put on feed. Two attempts were made to modify the early non-feeding experience of the chicks to see if this could account for the preference. In one case (Group *B,* Table 5) chicks were taken from a dark incubator, dyed a blue-green and were kept in a blue-green environment. In another case (Group *C,* Table 5) chicks were taken from the incubator and were placed in an environment in which the only illumination was through a green filter (Plexiglas Green 2092). In both cases (see Table 5) a preference was shown for the

TABLE 5

PREFERENCE OF CHICKS FOR A TAN OVER A DARKER BLUE-GREEN MASH, WHERE THE DIFFERENT GROUPS OF CHICKS HAD DIFFERENT KINDS OF EARLY VISUAL EXPERIENCE PRIOR TO THE TESTS

The number of pecks received by each color from *N* pairs of chicks is shown to the right of the notation for that color.

| | Group A | Group B | Group C |
|---|---|---|---|
| Tan | 586 | 164 | 147 |
| Blue-Green | 64 | 31 | 19 |
| *N* | 50 | 15 | 10 |

tan mash over the blue-green mash. The test apparatus used here consisted of the two kinds of mash fastened to the background, so that the chicks could peck at the mash but could not eat it.

The data of Table 2 suggest that feeding experience on the tan mash is not necessary in order for the chicks to show a preference for light desaturated colors. The data of Table 6 indicate that a preference for a light desaturated color can occur in spite of feeding on a dark different color.[3]

TABLE 6

THE EFFECT OF THE COLOR OF THE CHICK'S DIET ON THEIR PREFERENCE BEHAVIOR FOR MASH OF OTHER COLORS

Groups 8a and 9a were fed only on tan mash and Groups 8b and 9b on dark blue-green mash.

| Group 8a | $W_2$ | $R_4$ | $R_1$ | $R_4$ |
|---|---|---|---|---|
| No. of pecks | 69 | 20 | 119 | 21 |
| No. of pairs | 5 | | 6 | |
| Pecking ratio | 3.45 | | 5.67 | |
| Group 9a | | | | |
| No. of pecks | 100 | 23 | 461 | 51 |
| No. of pairs | 5 | | 7 | |
| Pecking ratio | 4.35 | | 9.04 | |
| Group 8b | | | | |
| No. of pecks | 338 | 87 | 662 | 187 |
| No. of pairs | 5 | | 7 | |
| Pecking ratio | 3.89 | | 3.54 | |
| Group 9b | | | | |
| No. of pecks | 301 | 94 | 508 | 102 |
| No. of pairs | 8 | | 8 | |
| Pecking ratio | 3.20 | | 4.98 | |

---

[3] The authors wish to thank Robert L. Fantz for collecting the data for this part of the experiment.

Here the chicks (Leghorns) were tested for their preferences by means of a display box similar to the one previously described except that four compartments were used instead of two. Each of two groups of chicks were divided into two smaller groups. Two of the smaller groups (one from each of the larger groups) were fed only on the tan mash and the others only on the dark blue-green mash (colored by dyeing the tan mash with methylene blue). The colored mash used in the tests was similar to that used previously. $R_1$ was slightly more saturated than the $R_1$ listed in Table 7 and $R_3$ was more saturated than the previous $R_2$. $R_4$ was unchanged. It will be seen (Table 6) that $W_2$ and $R_1$ were preferred over $R_4$ by the chicks which were fed only on the dark blue-green mash (Groups 8b and 9b) as well as by the chicks which were fed only on the tan mash (Groups 8a and 9a. In addition to the results in Table 6, 5 pairs of chicks were used in comparing $R_3$ and $R_4$. $R_3$ received 157 pecks and $R_4$ 115 pecks. The reinforcement on a low reflectance stimulus of one color did not result in the chicks preferring the low reflectance stimulus in the red series. It is probable, therefore, that experience on one color does not greatly influence the shape of the preference curve along a saturation series of another color.

TABLE 7
COLOR HARMONY MANUAL MATCH FOR THE COLORED MASH USED IN THIS EXPERIMENT

| Notation used in experiment | Color | Approximate manual value | Manual reflectance value |
|---|---|---|---|
| $W_1$ | white | b | 74 |
| $W_2$ | white | a | 88 |
| LG | light grey | d | 45 |
| MG | medium grey | f | 28 |
| DG | dark grey | l | 9 |
| $R_1$ | pale pink | 7 ca | 65 |
| $R_2$ | pink | 7½ ga | 39 |
| $R_4$ | red | 7½ pa | 12 |
| $B_2$ | light blue | 15 ga | 53 |
| $B_3$ | blue | 15 la | 30 |
| $G_2$ | light green | 22 ic | 37 |
| $G_3$ | green | 22 le | 22 |
| $G_4$ | green | 21 pi | 8 |
| Tan | tan | 2 lg-ie | approx. 25 |
| Blue-green | blue-green | 20 ng | 11 |

On the basis of these several kinds of evidence it is concluded that the pecking preference behavior of the chicks in this study is essentially natural (unlearned) preference behavior.

## E. SUMMARY AND CONCLUSIONS

Groups of chicks were tested for their color pecking preferences by using pairs of simultaneously presented mash stimuli. The results indicate that the chicks have a natural (unlearned) preference for light desaturated colors and that these preferences cannot be explained completely by the lightness of the stimuli.

## REFERENCES

1. COLOR HARMONY MANUAL, 3rd Edition. Chicago: Container Corporation of America, 1948.
2. HESS, E. H. Development of the Chick's response to light and shade cues of depth. *J. Com. & Physiol. Psych.*, 1950, **43**, 112-122.
3. LASHLEY, K. S. The color vision of birds. *J. Anim. Psychol.*, 1916, **6**, 1-26.
4. WALLS, G. L. The Vertebrate Eye and its Adaptive Radiation. Bloomfield Hills, Michigan: Cranbrook Inst. of Science, 1942.

*Department of Psychology*
*University of Chicago*
*Chicago 37, Illinois*

Published as a separate and in *The Journal of Psychology,* 1954, **38**, 495-502.

# PERCEPTUAL ADEQUACY IN A LIFE STRESS*[1]

*Institute for Psychosomatic and Psychiatric Research and Training, Michael Reese Hospital, Chicago*

SHELDON J. KORCHIN AND HAROLD BASOWITZ

## A. INTRODUCTION

There is little question that stress and anxiety have profound and far-reaching effects on many aspects of behavior. The present paper will report some findings on the effect of stress on perceptual functioning. In this experiment a simple perceptual function involving the discrimination of openings in circles presented tachistoscopically was studied as it varied during a prolonged period of life stress—paratroop training—and as it was related to a biochemical index of anxiety.

A sizable literature attests to the fact that both experimental stress and clinical anxiety can reduce the adequacy of perceptual functioning. Angyal (1) showed that anxious patients showed characteristically random behavior in the recognition of letter patterns presented tachistoscopically. Their errors primarily involved the substitution of wrong letters rather than the omission of correct ones. Postman and Bruner (10) similarly found that normal subjects under stress had a higher threshold for the correct perception of three-word sentences and were more reckless in their pre-recognition hypotheses. Experimentally stressed subjects were distinctly slower in completing and recognizing incompleted pictures (11). An experiment by Korchin, Singer, and Ballard (7) found a primitivation of the reproductions of geometric forms under conditions of failure-stress. These authors, as well as Postman and Bruner, suggest that the reduction of perceptual efficiency may be a special case of the more general primitivation of functioning proposed by Barker, Dembo, and Lewin (2) and demonstrated by them in the lowered constructiveness of children's play following frustration. In a later experiment Korchin (6) found that stressed subjects had a higher threshold for the recognition of slight openings in circles.

---

*Received in the Editorial Office on July 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Paper presented at the Sixty-First Annual Convention of the American Psychological Association, Cleveland, Ohio, September 5, 1953. This report is part of a larger study of anxiety supported by funds provided in U. S. Army Contract No. DA-49-007-MD-131 through the Medical Research and Development Board, Office of the Surgeon General, Department of the Army, Washington, D. C. A more complete account of the entire study will appear in a forthcoming book (4).

Another line of research has shown that patients with free anxiety characteristically have an elevation of hippuric acid excretion, to a level considerably above that of normals and other patients (9). The test consists of the intravenous administration of a standard dose of sodium benzoate and the measurement of the amount of hippuric acid (benzoyl glycine) excreted in a one-hour urine sample. It has conventionally been used as a measure of liver function and is known to estimate its ability to mobilize glycine. Ordinarily elevated in patients with early liver disease, anxiety patients show this hyperexcretion wholly in the absence of such disease. Not only is the level higher in patients with free anxiety, but it has further been found that hippuric acid varied with changes in anxiety, declining as anxiety lessened (8). On the assumption that hippuric acid excretion may define a scale of anxiety and stress proneness, the present experiment studies the sensitivity of extreme groups as they undergo the stresses of a life situation.

More specifically, this experiment was undertaken to test two principal hypotheses: (*a*) that perceptual functioning would be altered during a prolonged life stress; and (*b*) that proneness to such effects will be a function of initial hippuric acid level. Since the situation chosen involves a period of some duration it is further possible to study the changes in perceptual functioning as they may reflect the adaptation or cumulation of stress effects.

## B. Subjects

One hundred *S*s were randomly drawn from an incoming class at the Airborne Training Center at Fort Benning, Ga. Of this number the 10 highest and the 10 lowest in hippuric acid excretion were selected and formed into a single training group. Eight paratroop soldiers not in training and not jumping were used as controls.

## C. Experimental Situation

Training consists of three weeks during which men learn minutely the necessary skills of the airborne soldier. Each week has a focal and novel stress: during the first week the men jump each day from a 34-foot tower wearing a parachute harness connected to a pulley arrangement on which they ride down an inclined cable; in the second week they are lifted to the top of a 250-foot tower and dropped with open parachute; in the third week they are promoted to jumping from an aircraft in flight. But over and above the stress inherent in the jumps themselves, the entire atmosphere contributes to a generally high level of psychological tension. Discipline more rigorous than in the usual army unit, added to the trainee's intense desire to "win his wings" and preserve the respect of his classmates and the cadre make the fear of failure more acute and allow us to conceive the entire period as a continuous stress situation in which the jumps are only more intense foci.

## D. PROCEDURE

The experimental task (tachistoscopic closure test) consisted of 45 circles flashed by a projector fitted with a shutter onto a screen approximately nine feet away. Fifteen stimuli were presented at each of three exposure levels—1/100, 1/50, and 1/25 second. Of each set of 15, 3 were complete circles, and the remainder had 7° openings which appeared equally often at the top, bottom, right, and left of the circle. The task for *S* was to tell *whether* and *where* the circle was open. Prior to each day's trial *S*'s eyes were dark-adapted for 10 minutes with dark red goggles.

During training *S*s were tested immediately before the day's first jump. The "laboratory" was a tent within sight of the towers or planes. There were 17 trials in all. Three were on days prior to the start of training (base period); then *S*s were tested five times during the first week; on Monday, Wednesday, and Friday of the second week; four times during the third week; and twice after graduation. In addition, the test was repeated on Wednesday of each week after the end of the training day. On all trials the effort was made to keep conditions constant—the seating of the subjects, intensity of illumination, viewing angle, and so on.

## E. RESULTS AND DISCUSSION

Figure 1 presents the mean number of stimuli correctly perceived on each occasion by each group. These curves are based on the data for the 17 men who completed training of the original 20 *S*s. The high hippuric acid group (III H) consists of 8 *S*s, the low group (III L) of 9. The control group (VI), presented for comparison, includes 8 individuals tested on 11 occasions.

A number of findings are apparent on inspection. If the trainee groups, both the Highs and the Lows, are considered together and contrasted with the control subjects, it can be seen that the controls, not under the stress of paratroop training, do uniformly better on each of the paired occasions. It should be pointed out that the terms "base," "Week 1," and so on, refer to the training periods of the experimental *S*s; controls were tested on matched but not identical days. The differences between the training groups and control group, on each occasion separately, fall short of statistical sig-

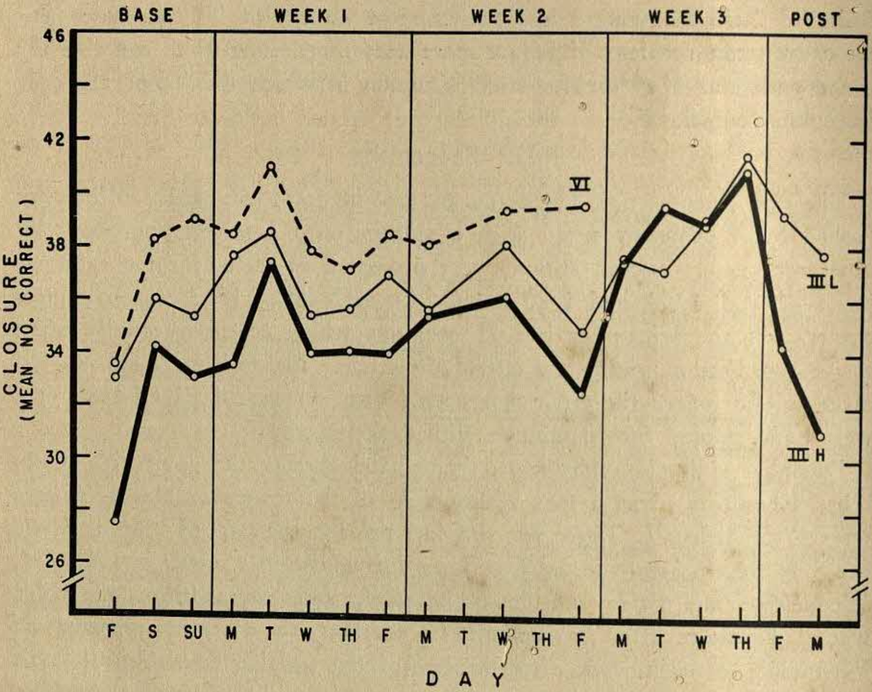


FIGURE 1
Tachistoscopic Closure Scores for High (III H) and Low (III L) Hippuric Acid Groups and Control Group (VI)

nificance. However, the consistency across all trials would seem to justify the conclusion that *S*s in training did indeed function more poorly.

Both the high and the low hippuric acid groups show essentially the same pattern of scores during training. For both there is a rise in score from the first to the fifth occasion, thereafter a plateau, then another rise during the third week. We do not find, therefore, a distinct difference between Highs and Lows in response to any part of the training situation. However, on 16 of the 17 occasions the low hippuric acid subjects are uniformly superior in performance. Although statistically the difference between the two groups, on any given trial, does not reach an acceptable level of significance, the implication is clear that *S*s selected as high in hippuric acid excretion are in fact more stress-prone and do function more poorly on this measure of perceptual adequacy. Thus we see that the stresses of airborne training do lower efficiency on a perceptual task, the more so in a group of *S*s predicted to be more vulnerable on the basis of a biochemical indicator.

On Wednesday of each week the closure test was readministered after the end of the training day. These measures may be considered as "post-stress" in the sense that they are removed psychologically from the focal point of the jump. This comparison reveals that during the first week there is a distinct rise in score from the pre-jump to post-stress measure and a drop back to original level the following morning (Table 1). All of these changes are statistically significant: for the Highs, pre to post, $t = 3.67$, $P < .02$; post to following morning, $t = 3.09$, $P < .05$; for the Lows, pre to post, $t = 3.54$, $P < .01$; and post-stress to following morning, $t = 3.85$, $P < .01$. During the second week the trend is the same although the rise from pre-jump to post-stress measures is not significant for either group. However, the drop from post-stress to following morning is about as great and is significant. For the Highs, $t = 3.66$, $P < .01$; for the Lows, $t = 3.32$, $P < .01$. In the third week, when the morning scores had already reached a fairly high level, there is no longer any post-stress improvement.

TABLE 1

COMPARISON OF MORNING (PRE-JUMP), EVENING (POST-STRESS) AND FOLLOWING MORNING (PRE-JUMP) MEAN CLOSURE TEST SCORES FOR HIGH AND LOW HIPPURIC ACID GROUPS ON WEDNESDAYS OF THE THREE TRAINING WEEKS*

| | Week 1 | | | Week 2 | | | Week 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | Pre-Jump | Post-Stress | Pre-Jump | Pre-Jump | Post-Stress | Pre-Jump | Pre-Jump | Post-Stress | Pre-Jump |
| Highs | 34.0 | 40.0 | 34.1 | 36.1 | 38.3 | 32.6 | 38.8 | 37.3 | 40.8 |
| Lows | 35.4 | 40.4 | 35.6 | 38.1 | 39.6 | 34.8 | 39.0 | 41.4 | 41.3 |

*The values in this table are the means for all the men tested on *each* occasion. The *t* tests reported in the text are based only on the scores of men tested on *both* of the occasions compared.

These data show perhaps more clearly than the differences between experimental and control groups the effects of stress on this function. When men are tested under more relaxed conditions—not under the anticipatory anxiety of having to jump—their performance is markedly improved, indeed equal or superior to that of the control group. After the training day, particularly during the first week, there is a definite reduction of tension and a consequent improvement in function.

Since the same *S*s were tested repeatedly we may examine the consistency with which individuals maintained their positions relative to the groups. How highly correlated is performance on different trials? The coefficient of concordance ($W$) for the high hippuric acid group is .47, significant at less than the .01 level; for the low hippuric acid group it is .68, again significant at less than the .01 level. Thus there is high intra-individual consistency in both groups, but the degree of consistency among the Highs is distinctly less than among the low *HA Ss*. Here again is an important difference between these two experimental groups. It seems as though the stress of airborne training not only lowers the overall level of performance, but contributes as well to an increased variability in the high *HA* group.

Interesting data are found in the post-training period. Despite the rise in mean performance in both the Highs and Lows during the third week of training, scores *drop* distinctly on the two days *after* training is completed (Figure 1). This decrement is so great that on the last day the Highs perform less well than they did on any previous day since the first. On first view these data are puzzling. Training is over, there are no jumps, and yet this measure which we believe sensitive to anxiety shows a distinct response. It might be claimed that motivation to do well on this task is reduced at this time. Now that the *S*s wear their wings, they try less hard, are less attentive, and do more poorly. Other evidence, however, negates this interpretation. Elsewhere (3) it has been shown that in the period following graduation these same *S*s have an increase in reported anxiety. A lowering of motivation might affect the closure test but it could hardly influence men to express increased levels of anxiety. We believe that the post-stress disturbance represents a genuine and important part of the stress response.

In order to function effectively throughout training men had to contain as rigidly as possible the potentially disruptive effects of anxiety. But when stress is reduced there is a corresponding release of control and rising anxiety with its concomitant psychological disturbance results. In the relative relaxation of the post-training period tension hitherto contained breaks through. This finding highlights a general characteristic of stress behavior. Given a stress which is not crippling in intensity, the individual will muster whatever resources are available in order to maintain adequate functioning. However, with the termination of the stress the controls and defenses invoked are relaxed, leading in turn to the appearance of anxiety and psychological disturbance. What was found in this experiment is, we believe, akin to the wartime observation that anxiety attacks often occurred following withdrawal from a combat zone (5).

The analyses presented thus far have considered only the total scores for each trial, i.e., the total number of stimuli correctly perceived. In these scores the *type of error* is of no concern. There are four types of error possible in this simple task: (*a*) an *open* circle may be seen as *closed* (closure error); (*b*) a *closed* circle seen as *open*; (*c*) an *open* circle seen as *open* but the position of the opening incorrectly located; and (*d*) the subject may not respond at all to a particular item. In all groups the greatest number of errors is of the closure type. However, both stress groups show a significantly *greater* proportion of closure errors than do the Controls. In the high *HA* group 79 per cent of the errors, and in the low *HA* group 81 per cent are of this type, compared to 64 per cent in the Control group.

Thus we find that not only do the stress subjects perform more poorly generally, but their errors seem to reflect less discrimination and a greater inability to resist the "pull of closure." This finding seems to us to have importance theoretically, since we conceive the anxious person as being characterized at least partially by cognitive dedifferentiation. In this state one would expect the ability to maintain the level of alertness and vigilance necessary to offset closure to be reduced.

## F. SUMMARY

Two groups of *S*s extreme in hippuric acid production were studied as they went through a life stress—paratroop training. Daily measures were made on a simple perceptual test prior to each day's jump. It is found that experimental *S*s under the stress of airborne training do show a reduction of perceptual efficiency. When anxiety is reduced during the evening of training days performance improves. At the end of training there is a distinct drop in performance which we have interpreted as resulting from a release of anxiety following the relaxation of controls. Proneness to these stress responses is reflected in the pre-stress hippuric acid excretion level, thus expanding further the relationship of this biochemical index to psychological anxiety.

## REFERENCES

1. ANGYAL, A. The diagnosis of neurotic traits by means of a new perceptual test. *J. of Psychol.,* 1948, **25**, 105-135.
2. BARKER, R., DEMBO, T., & LEWIN, K. Frustration and regression: an experiment with children. *Univ. Iowa Stud. Child Welf.,* 18, No. 1.
3. BASOWITZ, H., KORCHIN, S. J., & GRINKER, R. R. Anxiety in a life stress. *J. of Psychol.,* 1954, **38**, 503-510.
4. BASOWITZ, H., PERSKY, H., KORCHIN, S. J., & GRINKER, R. R. Anxiety and Stress. New York: Blakiston (to be published).
5. GRINKER, R. R., & SPIEGEL, J. P. Men Under Stress. Philadelphia: Blakiston, 1945.
6. KORCHIN, S. J. (Unpublished.)
7. KORCHIN, S. J., SINGER, J. L., & BALLARD, R. G. The effect of frustration on the reproduction of visually-perceived forms. *Personality,* 1951, **1**, 54-66.

8. PERSKY, H., GAMM, S., & GRINKER, R. R. Correlation of fluctuation of free anxiety and quantity of hippuric acid excretion. *Psychosomat. Med.*, 1952, **14**, 34-40.

9. PERSKY, H., GRINKER, R. R., & MIRSKY, I. A. The excretion of hippuric acid in subjects with free anxiety. *J. Clin. Invest.*, 1950, **29**, 110-114.

10. POSTMAN, L., & BRUNER, J. S. Perception under stress. *Psychol. Rev.*, 1948, **55**, 314-323.

11. VERVILLE, E. The effect of emotional and motivational sets on the perception of incomplete pictures. *J. Genet. Psychol.*, 1946, **69**, 133-145.

*Psychology Laboratory*
*Institute for Psychosomatic and Psychiatric Research and Training*
*Michael Reese Hospital*
*Chicago 16, Illinois*

Published as a separate and in *The Journal of Psychology,* 1954, **38,** 503-510.

# ANXIETY IN A LIFE STRESS*[1]

*Institute for Psychosomatic and Psychiatric Research and Training, Michael Reese Hospital, Chicago*

HAROLD BASOWITZ, SHELDON J. KORCHIN, AND ROY R. GRINKER

## A. INTRODUCTION

Free anxiety is a profound organismic state involving both somatic and psychological participation. In its psychological aspect it is *consciously experienced* and *subjectively reportable* as a feeling of intense dread and foreboding (5). The individual is beset by a generalized apprehensiveness; the dangers are objectless, and hence manifold in number. A host of stimuli and situations bear the capacity of activating or augmenting his condition of terror. This is anxiety in its purest form. Needless to say, the state may be found to exist in varying degrees among and within various groups, and within the same person at different periods in time.

Earlier research has demonstrated that a characteristic biochemical response in subjects having free anxiety is a distinct elevation of hippuric acid tolerance, well above the levels found in normals and other psychiatric patients (5). Hippuric acid production following a standard dose of sodium benzoate is a measure of liver function which reflects the rate of glycine mobilization by the liver for the conjugation of the benzoate. This test has been employed conventionally as an index of liver disease, diminished production of hippuric acid being coördinate with the severity of the disease. Hyperexcretion of hippuric acid, however, has otherwise been reported only in the earliest stages of liver pathology. Nevertheless none of the free anxiety subjects showing heightened levels of hippuric acid evidenced liver dysfunction according to a large number of additional clinical indices. Moreover, further study revealed that as anxiety was reduced, hippuric acid synthesis was lowered (4). The possibility exists, therefore, that this biochemical index provides a crude continuum of the magnitude of free anxiety. As part of a multidisciplinary program of research on anxiety, the present experiment attempted further de-

---

*Received in the Editorial Office on July 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Paper presented at the Sixty-First Annual Convention of the American Psychological Association, Cleveland, Ohio, September 5, 1953. This report is part of a larger study on anxiety supported by funds provided in U. S. Army Contract No. DA-49-007-MD-131 through the Medical Research and Development Board, Office of the Surgeon General, Department of the Army, Washington, D. C. A more extensive report of the entire study will be available in a forthcoming volume (1).

velopment of this relationship through the study of a *normal healthy* group of individuals undergoing *a real life stress* of some duration.

The life stress chosen was the paratroop training course at Fort Benning, Georgia. Classes consist of three weeks of graduated and intensive instruction in the requisite skills of an airborne soldier. Each of the weeks is highlighted by a focal stress. During Week 1 the trainee is required to jump daily from a mock tower 34 feet high. Week 2 is distinguished by free falls with open parachute from a 250-foot tower. The final week is devoted almost exclusively to jumps from an airplane in flight.

In addition to these particular events, the entire training period is conducive to a generally high level of psychological tension. The atmosphere is serious and business-like. Nobody walks, everyone runs as he executes his task and moves from station to station. Lapses of attention, slovenly performance, and breaches of good soldierly conduct swiftly result in having to "hit the dirt" to do 10 push-ups. Discipline is much more rigorous than that found in most army situations.

During the course of a pilot study in which paratroop candidates were interviewed both before and while in training, it became clear that the anxiety generated in this situation was not a simple unitary phenomenon. Conceptually, two types of anxiety were observable, associated with two rather distinct loci—shame and guilt. Despite our expectation that the primary source of anxiety should be related to the physical dangers of jumping, we found instead that feelings related to shame were often of greater importance. That is, the trainees were apprehensive, not so much about the severe threat to life and limb, but about the prospect of not measuring up to the task undertaken. Should one not "win his wings," he would be humbled before his peers and in the eyes of idealized older figures whom he was seeking to emulate. From the standpoint of theory, guilt anxiety is viewed as related psychodynamically to fear of destruction and death; shame anxiety as connected to fear of failure, humiliation, and the loss of self-esteem (6).

## B. SUBJECTS

In the research now being reported, 100 randomly selected candidates for the airborne course were tested for hippuric acid excretion prior to training. All of the values were found to lie in the "normal" range, although both extremes were represented. The 10 highest and 10 lowest individuals in hippuric acid tolerance were then chosen for the experiment, yielding two groups with no overlap in terms of initial level of measurement. Unlike the previous studies in which anxiety was the independent variable, hippuric acid now became the independent variable.

## C. HYPOTHESES

Two hypotheses were posed: (*a*) that the "high" hippuric acid group (III H) would manifest greater susceptibility to anxiety under the stress of training than would the "low" hippuric acid group (III L); and (*b*) that the anxiety concerning fear of injury (guilt) would differentiate these two groups more sharply than would anxiety concerning fear of failure (shame), since guilt anxiety, so dramatically exemplified in the traumatic neuroses of war, is conceived as the more disorganizing of behavior.

## D. PROCEDURE

The method of measurement was quite simple. In response to the question "How do you feel about your chances of being hurt (failing) in airborne training?", subjects were required to rate themselves separately on (*a*) fear of injury, and (*b*) fear of failure on each of three days immediately before, almost daily during the three weeks of training, and twice after graduation. Both scales were marked off in units of 1, ranging from 0 to 6, with the extremes defined as "completely at ease and unafraid" at the 0 point, and "scared shitless of being hurt (failing)" at the opposite pole. The trainee merely circled the number appropriate to his feeling state on the particular day of measurement. Testing took place very early in the morning. During training, it occurred in the field just prior to the beginning of instruction so as to capitalize on the anxiety anticipatory to the day's activities.

## E. RESULTS AND DISCUSSION

The findings presented are based only upon *S*s who completed the airborne course. Attrition of training reduced the high and low hippuric acid groups to *N*s of 8 and 9 respectively. Figure 1 describes the mean ratings across all occasions for the two groups on the "harm" scale. It is quite apparent that the first four occasions of measurement, which include the base or pre-training period and the first day of training, reveal no appreciable differences between groups on any given occasion. Thereafter, however, throughout the entire training period (Tuesday, Week 1 through Thursday, Week 3) the groups diverge somewhat sharply as the Highs (III H) consistently rate themselves higher than the Lows (III L) on fear of harm.

It is perhaps noteworthy that while Monday of Week 1 shows a lack of distinction between the two groups, it is in fact a before-training measure in the sense that the ratings were made before the formal instruction had begun. Quite possibly some prior experience with the hazards of training is necessary before one knows of what to be afraid and how much.

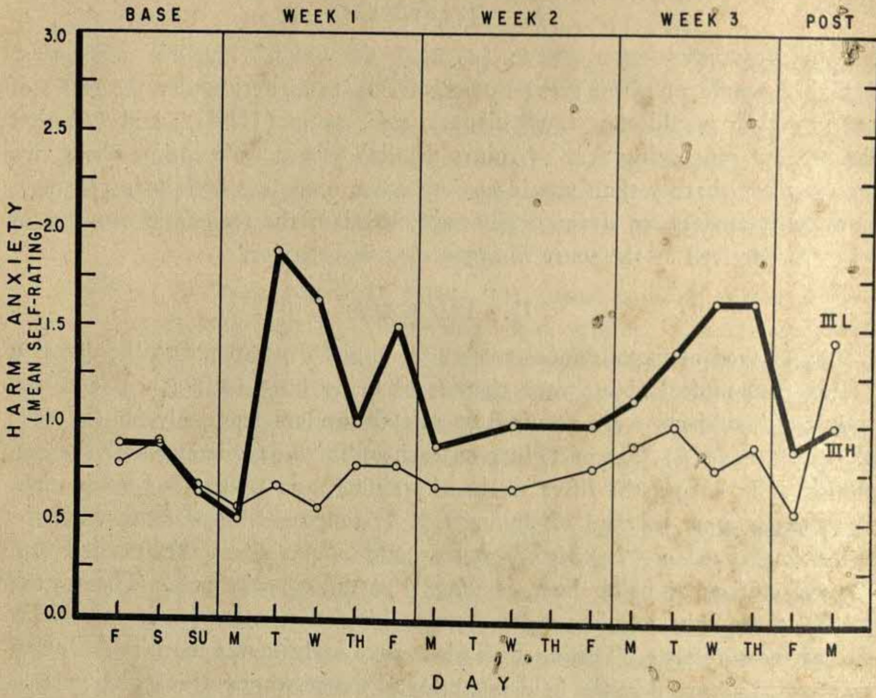


FIGURE 1
DAILY SELF-RATINGS ON HARM ANXIETY FOR HIGH (III H) AND LOW (III L) HIPPURIC ACID GROUPS BEFORE, DURING AND AFTER AIRBORNE TRAINING

When *t* tests are applied to estimate the significance of differences between means on any single trial, they fall short of the conventional .05 level. Thus the distributions clearly involve some degree of overlap between III H and III L on like days. Nevertheless, comparison of the entire training phase with the "base" measures tends to support the general proposition that, under the stresses of airborne training, individuals initially high in hippuric acid are more vulnerable to apprehensiveness concerning injury.

The post period contains certain curious phenomena. As was expected, the morning following the last airplane jump (Friday) finds both groups declining in reported anxiety relative to their just previous levels. However, three days after (Monday) they show a noticeable rise, especially in the case of the low hippuric acid *S*s who peak to their highest mean value for the entire experiment. This increase in anxiety following graduation has been termed the "end-phenomenon." There is firm reason to assume that it is neither accident nor artifact, since ample confirmatory evidence for its realness exists in perceptual and biochemical measures taken independently the same day (1, 2).

In accounting for its occurrence, we propose that the individual must exercise a maximum of control in order to cope with the many threats inherent in the training situation. With training ended, such control is loosened, permitting the eruption of anxiety contained heretofore. This is not unlike the instances reported during World War II when troops rotated from the front lines or combat air missions developed severe anxiety after they had reached rest areas removed from hazard (3). Just why III L should be elevated so strikingly on the last occasion is puzzling. No explanation is entirely satisfactory. Perhaps the greater compression of anxiety achieved by this group in the earlier phases results in an intensified swing when the checks are relaxed.

Still another characteristic of the curves in Figure 1 deserves comment. Inspection reveals that the variability of means for the Highs easily exceeds that observable for the Lows *during training.* A more quantitative estimate is achieved in the comparison of distributions of absolute change scores between adjacent points, e.g., Monday-Tuesday, Tuesday-Wednesday, etc. A test for the significance of differences between the means of arrays so prepared shows III H to be more variable than III L at approximately the .05 level of confidence ($t = 2.22$, $df = 10$). Thus, not only does the group initially high in hippuric acid tend to develop a generally higher level of anxiety concerning injury, but it is more variable in self-reported feeling state as well.

The difference in the trends of the two curves is a matter of further interest. Attrition rates in the Airborne Department plainly indicate that Week 3 poses the most and Week 1 the least physical danger throughout training. All paratrooper students are well indoctrinated with this information. Hence if ratings were to correspond to this "reality," the result should be a linearly rising curve. The line of best fit for the Lows approximates such a form; that for the Highs does not. The latter evidence exaggerated harm scores in a period when there is less objective reason for dread of injury—Week 1. In this sense they are less "realistic." This problem will be treated more fully below.

Figure 2 presents the data related to anxiety of failure. In general the findings parallel those encountered in the previous scale. Both groups have similar means for each of the first four occasions. With the onset of training, however, high hippuric acid *S*s tend to exceed the Lows on comparable days. Yet several reversals obtain and the trend, although suggestive, is not as marked as in the case of the harm scale. The "end-phenomenon" of increased

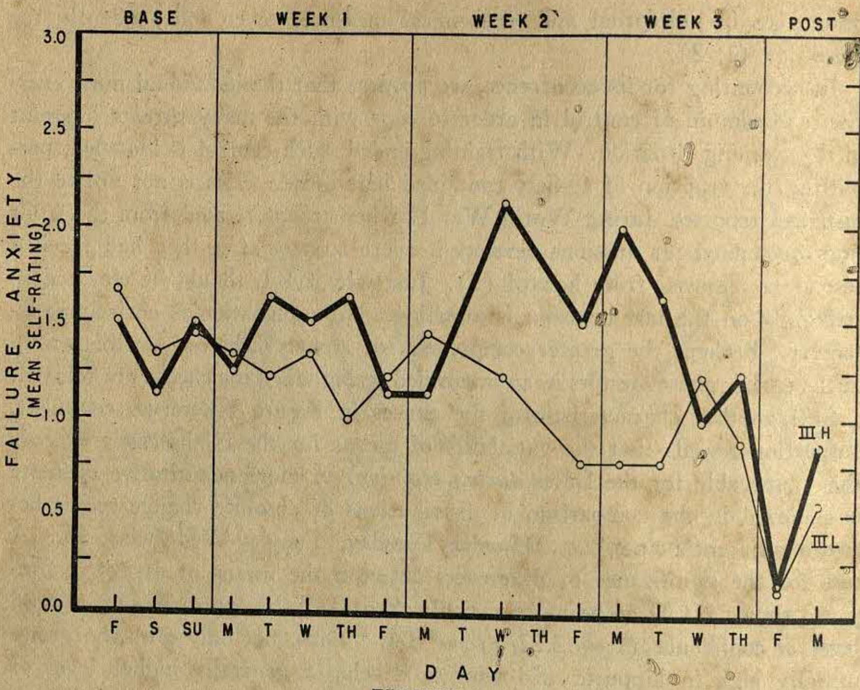


FIGURE 2
DAILY SELF-RATINGS ON FAILURE ANXIETY FOR HIGH (III H) AND LOW (III L) HIPPURIC ACID GROUPS BEFORE, DURING AND AFTER AIRBORNE TRAINING

anxiety three days after graduation (Monday, post) again appears. The same explanation is invoked: removal of controls is followed by a release of anxiety. Greater variability also is seen to typify the Highs during training. Statistical comparison of absolute change scores for the two groups yields a $p < .10$ and $> .05$ ($t = 2.10$, $df = 10$).

Airborne Department records reveal that the overwhelming proportion of candidates fail the course in Week 1 (approximately 85 per cent of all failures). By contrast the smallest number of failures occurs during Week 3. A theoretical curve consistent with these facts should be linearly decreasing. The line of best fit for III L approaches such a functon. On the other hand, the curve obtained by III H is characterized by heightened fear of failure at *less appropriate* times—from the middle of Week 2 to the middle of Week 3.

These observations may perhaps be clarified in still another manner. Since the *expected* curve for harm scores is opposite to that *expected* for failure scores (ascending vs. descending functions), self-ratings on the two scales should be negatively correlated—if feeling state is commensurate with the realistic dangers of the situation. The means of III L paired by day yield a rho of —.75, which is highly significant ($p < .01$) and in the predicted direction. The rho for III H is —.05, reflecting the more haphazard pairing of mean scores on the scales. There is here a provocative hint of greater anxiety being manifest through lack of correspondence between what a person fears and what objectively there is to fear. The implications of this point to the opening statements that dangers are objectless in the state of anxiety should be quite clear.

## F. Summary and Conclusions

In sum, it appears that the two hypotheses with which this experiment was undertaken are upheld. (*a*) Persons initially elevated in hippuric acid tolerance develop higher levels of anxiety under the stresses of airborne training than do those initially low on this measure, although prior to training the groups are indistinguishable in their self-ratings of anxiety. (*b*) Anxiety of injury (guilt) discriminates the two groups more clearly than does anxiety of failure (shame). Further, high hippuric acid subjects are characterized by greater variability in self-reported feeling state and achieve increased levels of anxiety at points in training less appropriate to the objective realities of the situation.

Thus the present study suggests an elaboration of the relationships found earlier between hippuric acid and free anxiety. Not only does hippuric acid excretion distinguish neurotics with free anxiety from normals and other patients, but it differentiates among normals as well. Clinically normal individuals whose excretion values are very high, although still below levels typifying the condition of free anxiety, are prone toward heightened anxiety in a life stress. Additional research could profitably be directed towards corroborating these findings on larger series of subjects and testing more precisely the predictive power of the hippuric acid index.

## References

1. Basowitz, H., Persky, H., Korchin, S. J., & Grinker, R. R. Anxiety and Stress. New York: Blakiston (to be published).
2. Korchin, S. J., & Basowitz, H. Perceptual adequacy in a life stress. *J. of Psychol.,* 1954, **38**, 495-502.
3. Grinker, R. R., & Spiegel, J. P. Men Under Stress. Philadelphia: Blakiston, 1945. Pp. 484.
4. Persky, H., Gamm, S. R., & Grinker, R. R. Correlation of fluctuation of free anxiety and quantity of hippuric acid excretion. *Psychosomat. Med.,* 1952, **14**, 34-40.